Minnesota's Statewide Quality Reporting and
Incentive Payment Program:
Final Recommendations to the Minnesota
Department of Health

March 25, 2009

Introduction

In 2008, the State of Minnesota through the Minnesota Department of Health (MDH) contracted with Minnesota Community Measurement (MNCM) for assistance in building and implementing a unified statewide quality reporting system for health care providers, and in developing a quality incentive payment system. This report is the final in a series of reports, and includes:

- Specific recommendations for the ongoing development and alignment of quality measures to be used for reporting for physicians and hospitals in Minnesota.
- Recommendations for a methodology to be used in developing an initial quality incentive payment system for physicians and hospitals, applicable to the State Employee Group Insurance Program (SEGIP) and the MN Department of Human Services (if Federal Law allows).
- An initial subset of uniform quality measures to be used in the first year of the quality incentive payment program (2010).
- A summary of the process and community stakeholders involved in defining both the quality measures and the quality incentive payment system.

Background

Through MN Community Measurement (MNCM), the Minnesota health care community has pioneered collaborative health care reporting: building a set of measures that have become both more sophisticated and less administratively burdensome; establishing a process that allows for the collection of quality measure data from medical groups as well as health plans; and providing for the reporting of Minnesota quality data to health care providers and to consumers.   In 2008, MN Community Measurement was contracted by the Minnesota Department of Health (MDH) to build a unified statewide quality reporting system for health care providers, to expand the number of quality measures for public reporting, and increase the number of physician clinics reporting data. Other goals of the contract include broadening the stakeholders involved in the measurement process, and enhancing consumer access and understanding of the quality measures. MN Community Measurement was also tasked to develop and implement a Quality Incentive Payment System for both physicians and hospitals across the State of Minnesota.

MN Community Measurement worked with a number of community partners including Stratis Health, the Minnesota Medical Association, the Minnesota Hospital Association, the Minnesota Council of Health Plans, and the University of Minnesota School of Public Health to complete various tasks and

reports for this initiative. Numerous public meetings were held across the state to seek community input and feedback, and commentary was solicited via a public e-mail address. A Hospital Quality Reporting Steering Committee comprised of representatives from both rural and urban hospitals, health plans, and purchaser/consumer representatives was formed to assist MNCM in the recommendation of a subset of quality measures for payment incentives for hospitals.   A separate Payment Incentive Workgroup including representatives from health plans, purchasers, and providers, was utilized to offer expert feedback on the physician quality measures and the overall methodology of the Quality Incentive Payment System. MNCM will continue working with these groups as we recommend future measures for reporting and incentive payment. All comments and feedback received during the process will be available in separate, supporting documentation.

Previous work completed under the MDH contract is attached in Appendices to this final report, and includes the following:

- In December 2008, the University of Minnesota, under subcontract with MNCM, completed an inventory of measures in use across the country for public reporting of quality information. The University's work was augmented with information from Stratis Health, the MHA, and the MMA. (Appendix A)
- In February 2009, Minnesota Community Measurement, along with community partners including Stratis Health, the Minnesota Medical Association, the Minnesota Hospital Association, and the University of Minnesota, recommended a set of quality measures for public reporting for the State of Minnesota for both clinics and hospitals. (Appendix B)
- In February 2009, the University of Minnesota completed Preliminary Recommendations of Measures and Methodologies for Minnesota's Quality Incentive Program. Included with the report was an Inventory and Literature Review of Pay-for-Performance Methods and Structure; and an Inventory of Performance Measures in Current Use for Pay-for-Performance Programs. (Appendices C,D,E)

This final report addresses the measures selected and the methodologies developed for the Quality Incentive Payment System, which will be implemented by July 1, 2009.

Measure Development and Selection

The University of Minnesota's Preliminary Recommendations emphasized that measure selection for public reporting and incentive payments will be an iterative process as new science and updated methodologies will result in revisions to existing measures and/or replacement of measures with newly developed measures. We heard from stakeholders that they recognize the timeframe for measure

development, which involves collaboration with providers and other payers, and testing or piloting measures before they become standardized reporting measures. Both providers and health plans agree on the need for continued development of new measures, especially Specialty Care measures and measures that cross clinic and hospital settings to reflect patient experiences across the continuum of care. We heard general agreement that alignment of these measures across the community is also advantageous, but should not hamper innovation. A community measure alignment process which seeks to expedite alignment, yet allow for innovation is included in these recommendations. (see Diagram in Appendix F)

Table 1 presents the general timeline for the development and implementation of measures for public reporting through 2010. Measures used for quality incentives will be selected from this process.

Table 1. The measures shown are the minimum number of measures that will be added to the statewide quality reporting system in each year of the contract. The different colors are used to show how a measure works its way through the multi-stage development process over time.

Measurement Development Process Stages

	Stage 1: Determine and select future measurement priorities; Develop new measures	Stage 2: Data collection on new measures begins; Voluntary data submission; Voluntary public reporting	Stage 3: Public Reporting Statewide
First Year: 2009 Calendar Year	1 Primary care measure 1 Specialty care measure 5 New hospital measures supported by clinical-data enhanced database[1]	Depression measure – primary care Depression measure – specialty care Health information technology (HIT) measures Patient experience measures	Existing MNCM and Minnesota Hospital Quality Report[2] measures 12 Additional AHRQ[3] inpatient hospital measures
Second Year: 2010 Calendar Year	1 Additional primary care measure 2 Additional specialty care measures Continuation from first year of development of 5 new hospital measures supported by clinical-data enhanced database[1]	1 Primary care measure 1 Specialty care measure	Previous year's measures, plus: Depression measure – primary care Depression measure – specialty care HIT measures Patient experience measures 2 Additional AHRQ[3] inpatient hospital measures
Third Year: 2011 Calendar Year	2 Additional specialty care measures	1 Additional primary care measure 2 Additional specialty care measures	Previous year's measures, plus: 1 Primary care measure 1 Specialty care measure 5 New hospital measures supported by clinical-data enhanced database[1] 2 Additional AHRQ[3] inpatient hospital measures
Fourth Year: January 1, 2012 – July 1, 2012		2 Additional specialty care measures	Previous year's measures, plus: 1 Additional primary care measure 2 Additional specialty care measures 2 Additional AHRQ[3] inpatient hospital measures

[1] Clinical-data enhanced database will integrate clinical data with administrative data; [2] The Minnesota Hospital Quality Report can be seen at www.mn.hospitalquality.org; [3] AHRQ is the Agency for Healthcare Research and Quality

The MNCM Reporting Advisory Committee (RAC), which includes physicians and other clinicians, purchasers, consumers, technical specialists and health plans, will approve all ambulatory measures for the statewide quality reporting system. The MNCM Board of Directors will also approve all ambulatory and hospital measures prior to implementation.

According to the enabling legislation, in selecting a subset of the reporting measures for the Quality Payment Incentive system, the focus should be placed on outcome-related measures that improve care and lower costs for high volumes of people; address chronic conditions; and minimize both providers' administrative burden and duplication of related activities. We recommend that the initial phase of the Incentive system should focus on the pay-for-performance methodology and not add to the implementation burden by introducing new performance measures.   The initial measures selected are well-established and are already being utilized by other payers in the state and/or by other national payers.

As incentive measures for dates of service in <u>calendar year 2009</u> (reported in 2010) we recommend:

Ambulatory Care Measures (collected directly from providers):

- Comprehensive Diabetes Care (Institute for Clinical Systems Improvement (ICSI) Health Care Guideline)

- Optimal Vascular Care (ICSI Health Care Guideline; HEDIS 2009 Cholesterol Management for Patients with Cardiovascular Conditions)

- Depression PHQ-9 Six month remission (ICSI Depression Action Workgroup; DIAMOND Project)

MNCM will continue to recommend ambulatory incentive measures on an annual basis. A summary of the selected measures (with guideline citation), data collection processes, and reporting timelines will be available to providers well in advance of reporting.

Hospital Care Measures (reported by providers and externally validated):
For both Prospective Payment System (PPS) hospitals and Critical Access Hospitals (CAHs), the following Hospital Compare Measures - Appropriate Care Measures* (Centers for Medicare & Medicaid Services (CMS):

- AMI/heart attack

- Heart Failure

- Pneumonia

*The Appropriate Care Measure is a patient-centered all-or-none approach to measurement that recognizes when a patient receives all of the evidence-based care for which they were eligible, and gives no 'credit' when all care is not given.

In addition to the three measures from 2009, the following hospital measures are recommended for dates of service in 2010 (reported in 2011) for Prospective Payment System (PPS) hospitals, and Critical Access Hospitals (CAHs) if deemed relevant after one year of data review:

- Hospital Compare Measure: A new composite measure for Surgical Care Improvement Project (SCIP), which includes the core SCIP measures as currently reported on the Minnesota Hospital Quality Report and the following two measures from the Hospital Acquired Infection Measures group mandated by the Minnesota legislature:

    i. Cardiac surgery patients with controlled 6 a.m. postoperative blood glucose
    ii. Surgery patients with appropriate hair removal

- AHRQ Measures:
    o Decubitus ulcer
    o Death among surgical patients with treatable serious complications
    o Post-operative pulmonary embolism or deep vein thrombosis
    o Obstetric trauma – vaginal delivery with instrument
    o Obstetric trauma – vaginal delivery without instrument

- A newly developed measure of patient experience derived from the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey

Methodology for Quality Incentive Payment System

Our recommendations on methodology for the Quality Incentive Payment system were designed for initial implementation by the State Employees Group Insurance Plan (SEGIP) and the MN Department of Human Services (if allowed by federal law). We expect the model to be adjusted and refined in future years. We recommend an initial model that includes rewarding providers who attain a defined benchmark performance level, or a defined improvement goal designed to close the gap between prior performance and 100%. We recommend setting annual benchmark levels by reviewing historical data

and calculating a stretch goal designed to reward the top 20% of providers. The 2009 benchmark levels and improvement goals will be determined early in the second quarter of 2009, pending the final analysis of data from 2008. Annual benchmarks and improvement goals should be posted to the MDH website, along with a summary of selected measures, well in advance of the reporting process.

We recommend that ambulatory measures be risk adjusted through manipulation of the specific inclusion and exclusion criteria of measure specifications, combined with an adjustment of provider case mix by major product types – Medicaid, Medicare, Uninsured, and Commercial populations. The hospital risk adjustment methodology should have a continued focus on co-morbidity and severity. In future years, we recommend an evaluation of additional risk adjustment models, and consideration of a more complex measure weighting system.

Recommendations

The following recommendations are the result of a collaborative process involving input and feedback from health care providers, payers, and consumers from across Minnesota. Initial recommendations provided by the University of Minnesota were used as a starting point for discussions. General agreement was reached on the following final recommendations. Stakeholders voiced the need for a system that fosters the continued sharing of best practices and collaborating for improvement. In addition, measures should focus on areas in need of systemic improvement and central to transformational change.

Note: Recommendations related to the development and implementation of a Payment Incentive System (#5 through #11), are relevant specifically to the State Employees Group Insurance Plan (SEGIP) and the MN Department of Human Services (if allowed by Federal law).

General Recommendations

Recommendation 1
All of the recommendations put forth in this document should apply to Minnesota health care providers (both Ambulatory and Hospital), as defined in section 62J.03, subdivision 8 of Minnesota Statutes, at the site of care level.

Rationale:

Historical measurement experience has shown wide variation in results among individual clinics or hospitals, within the same medical group or system. Measuring at the individual site level provides increased incentives to providers for improvement, and additional transparency for consumers.

Comments:

All parties agreed.

Recommendations Related to Measures

Recommendation 2

The Statewide Standardized Quality Measures referred to in Minnesota Statute 62U.02 should include:

- **Measures approved for the Statewide Quality Reporting System and publicly reported by MNCM, as set forth in the February 9, 2009 Report.**
- **Other measures not publicly reported by MNCM that:**
    - ○ **Measure the results of care for conditions not otherwise measured in the Statewide Quality Reporting System, and**
    - ○ **Are aligned across all health plans, and**
    - ○ **Have been reviewed and approved by MNCM Reporting Advisory Committee and Board of Directors as valid, reliable and useful.**
- **Measures approved for HEDIS reporting by Health Plans.**
- **Measures required for eValue8 reporting.**
- **Other measures necessary for compliance with Federal or State laws.**

Rationale:

The Statewide Standardized Quality Measures system should balance the need for alignment of measures with the need to rapidly develop and test new measures. Provider organizations expressed concern that any measures that require all providers to participate should include a community wide review process. Health plans expressed concern that a lengthy review process could increase the time needed to implement new measures that would be valuable to patients and the community. This recommendation would allow alignment of measures without restricting the development and testing of new measures on a voluntary basis between health plans and provider organizations.

Comments:

Additional clarification of the Minnesota law was requested and discussed by all parties.

Recommendation 3

Use well established performance measures for introducing a statewide program of pay-for-performance.

An initial list of Statewide Standardized Quality Measures for public reporting was recommended in a February 6, 2009 report to MDH. All measures on the list have been formally approved by MNCM's Report Advisory Committee and Board of Directors, and will be available to all payers for use in incentive payment or pay-for-performance programs.

Rationale:

Large scale and community-wide examples of pay-for-performance usually followed or were associated with measures that had already been used for public reporting or the private profiling of providers, or piloted in "dry runs". We think a statewide initiative, such as this, requires that the initial measures be well-established in the community. Measures of physician practice already implemented by MNCM and supported by health plans, as well as measures already implemented by Minnesota hospitals for public reporting, should be regarded as the logical initial candidates for pay-for-performance, with other measures added after they have been piloted.

Comments:

There was general agreement that the established measures we have been reporting in Minnesota are appropriate for state-wide payment incentives. Health plans expressed some concerns about whether this recommendation would restrict their ability to test new measures with payment incentives.

Recommendation 4

While initially using a subset of the Standardized Quality Measures and Hospital Compare measures for incentive payment, the State should develop a plan to increase the use of clinical outcome measures in the Standardized set, including patient experience measures.   Efforts should be made to expedite the development of Specialty Care measures, and measures that cross clinic and hospital settings.   Additional measure development and innovation should continue and be encouraged

among all payers in the State.   Measure alignment should continue across health plans, using a collaborative Community Alignment Process (see diagram - Appendix F):

- If a health plan implements a new quality measure as a contractually required reporting measure for providers, the alignment process will be initiated to determine one consistent measure across all health plans. Aligned measures will be approved by MNCM's Report Advisory Committee and the Board of Directors, and approved measures will be listed in a Standardized Quality Measures Catalogue. This Catalogue will be updated at least annually.
- Quality measures used in short-term "Pilot" projects will be exempt from this alignment process for a period of one to two years. A Pilot measure is defined as:
  - A measure that does not have system-wide implementation; and
  - It has been agreed to by the health plan and the provider; and
  - The agreement includes a negotiated dollar award.

  Pilot measures may voluntarily be brought forward to become standardized measures, and will then be subject to the Community Alignment Process. If approved, the measure will be added to the Standardized Quality Measures Catalogue and become available to all health plans.
- Measures required for prior authorization of payment are exempt from this process.

Rationale:

The principle of using well established measures limits the number of possible different measures available for the initial phase of implementation, but experimentation with new measures, and with redesign of existing measures, should be ongoing as a stimulus to performance improvement. The alignment of measures alleviates the reporting burden on providers, but the process allows piloting of new measures.

Comments:

This alignment process is a compromise between health plans who voiced concern over retaining the ability to innovate, and providers who are seeking greater alignment of measures across all payers in the system. Health plans see the community alignment measure approval process as too lengthy, and prohibitive in launching new measures, especially important Specialty measures. Providers felt that continued measure development outside of the alignment process may result in health plans circumventing the intent of the law, but agreed that adding new Specialty measures was very important.

Recommendation 5

For the incentive plan used by SEGIP in 2010, begin with a relatively small subset of MNCM and Hospital Compare measures that are already use.

5a. Ambulatory measures recommended for 2010 SEGIP Incentive Program (Reporting 2009 Dates of Service) are:

Recommended 2010 SEGIP Incentive Program Measures:
- *Optimal Diabetes Care*
- *Optimal Vascular Care*
- *Depression*
- *AMI/Heart Attack Appropriate Care*
- *Heart Failure Appropriate Care*
- *Pneumonia Appropriate Care*

1. Optimal Diabetes Care - the percentage of patients with diabetes (Types 1 and 2) ages 18-75 who reached all five treatment goals:

    a. HbA1c <8

    b. Blood Pressure <130/80

    c. Low Density Lipoprotein (LDL) <100

    d. Daily Aspirin Use

    e. Documented Tobacco Free

2. Optimal Vascular Care - the percentage of patients with vascular disease ages 18-75 who reached all four treatment goals:

    a. Blood Pressure < 130/80

    b. LDL <100

    c. Daily Aspirin Use

    d. Documented Tobacco Free

3. Depression measure, primary care –

    a. Six Month Remission Rate (PHQ-9 score <5 at six month follow-up)

We further recommend that the State's formal rule regarding selected measures not include full and detailed measure specifications of the approved measures, but instead contain a summary of the measures selected annually.

Rationale:

If recommendation 3 is adopted, the field of candidate measures is relatively limited for the first phase of implementation. These three initial measures are already being used for incentive programs by major payers in the community, and are well-established community improvement initiatives. In addition, there is little evidence to date that focusing rewards on a small set of measures leads to poorer quality in areas not eligible for rewards. Expanding the number of measures over time should be a relatively short-

term programmatic objective, but scaling up a comprehensive value-based purchasing structure will require addressing most clinical care systems and processes in some way.   We also recommend that the State not include "overuse" performance measures, such as an antibiotic use for URIs, in its initial pay-for-performance program. It may be controversial to establish a state-sponsored pay for performance program that may appear to pay for withholding a service.   However, these measures can have important impact, and should be added in subsequent years of the program.   Finally, as measurement specifications can change due to changes in science, it is important that measures can be revised quickly. A formal rule changing process may delay revisions to important measures used in the system that impact appropriate patient care.

Comments:

Several payers felt strongly that "overuse" measures be available for reporting and incentive programs and should be considered for use by SEGIP in future years. Providers were in general agreement.   All parties voiced concern around including measurement specifications in the rule process, as this would slow down development and hamper the ability to change or update an existing measure that may affect patient care.

5b.   Hospital measures recommended for 2010 SEGIP Incentive Program (Reporting 2009 Dates of Service) for Prospective Payment System (PPS) hospitals and Critical Access Hospitals (CAHs).

Hospital Compare Measures:

1. AMI/heart attack: Appropriate Care Measure
2. Heart Failure: Appropriate Care Measure
3. Pneumonia: Appropriate Care Measure

For 2010 dates of service (reporting 2011), in addition to the 2009 measures above, we recommend the following measures in the incentive payment program for Prospective Payment System (PPS) hospitals, and Critical Access Hospitals (CAHs) if deemed relevant after one year of data review:

1. Hospital Compare Measure
    a. A new composite measure (to be developed in 2009) for Surgical Care Improvement Project (SCIP) that includes the core SCIP measures as currently reported on the Minnesota Hospital Quality Report and the following two measures from the Hospital Acquired Infection Measures group mandated by the Minnesota legislature:
        i. Cardiac surgery patients with controlled 6 a.m. postoperative blood glucose

<ol type="i" start="2">
Surgery patients with appropriate hair removal

2. AHRQ Measures

   a. Decubitus ulcer

   b. Death among surgical patients with treatable serious complications

   c. Post-operative pulmonary embolism or deep vein thrombosis

   d. Obstetric trauma – vaginal delivery with instrument

   e. Obstetric trauma – vaginal delivery without instrument

3. A newly developed measure of patient experience derived from the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey

Rationale:

While some of the individual Hospital Compare measures in AMI, heart failure, pneumonia, and surgical care are reaching quite high levels of performance across the state, there is still ample room for improvement when looking at the Appropriate Care Measure for each of the topics specifically for heart failure and pneumonia. In addition, heart failure and pneumonia in particular are relevant for Critical Access Hospitals (CAHs), and the calculation of the Appropriate Care Measure helps to address the low volume in some CAHs by aggregating all patients into one measure. For methodological reasons related to the types of surgeries included in the SCIP measure, it is inappropriate to calculate an all-or-none Appropriate Care Measure for SCIP. However, a composite weighted score should be developed, incorporating several measures from the Hospital Acquired Infection measures. Regarding the AHRQ measures, Critical Access Hospital data should be evaluated for usefulness and relevance in year one of public reporting, then consider if some or all of the 5 measures listed are appropriate for incentive payment for CAHs in year two. Finally, patient experience measures are relevant for hospitals of all sizes and types, but it is important to capture the correct questions from HCAHPS survey results. Although "overuse" measures are very important to health care, they may provide unintended consequences in a payment incentive program. These measures are not appropriate for initial implementation, but are urgently and strongly supported immediately thereafter, once a well defined hospital measures can be developed. The advancement of Health Information Technology may provide additional ways to measure overuse.

Comments:

The workgroup placed emphasis on measuring both PPS Hospitals and CAH Hospitals, and had a strong interest in transformational change of the hospital system. Multiple parties expressed interest in adding cost and efficiency measures.

Recommendations related to Methodology

Recommendation 6

Implement an incentive payment model that includes a combination of pay for improvement and pay for attainment. Encourage alignment with this model across all payers, including alignment of the benchmarks chosen for payment. For the SEGIP incentive program:

Reward for attainment of health care goals or for improvement towards those goals.

- The attainment or benchmark for each measure will be established annually, by reviewing historical data for each measure, calculating a stretch goal of 4-5% average improvement over the previous year, and targeting the top 20% of clinics reporting that measure.
- The annual improvement target will be a 10% reduction in the gap between the prior year's result for that provider and 100% (or another appropriate target goal for the population). A three year rolling average of the reduction in the gap will be calculated for each provider, and will be used to determine a reward. The most recent year of the three year average will be weighted more heavily. The three year average must be at least 10% of the gap between the prior year's result and 100% (or target goal) to attain a performance reward.
- Providers can receive the payment for either achieving the attainment or the improvement benchmark, but not both.

Rationale:

This two part system should be fine-tuned over time to ensure that providers continue to have an incentive to improve even after reaching the minimum target level. This approach must be modeled so that the payers' budget management requirements and the desire for the incentive to have an impact on provider behavior are considered in the context of current performance levels and variation in performance in Minnesota. Since alignment may have a bigger impact than specific methodology, SEGIP should have some flexibility in implementation of the methodology to allow alignment with other payers not included in this legislation, as long as they are consistent with the general recommendations.

Comments:

Some participants voiced concerns regarding how the combination of attainment and improvement would actually be weighted in determining rewards, and if there would be a meaningful return on investment.   Others agreed this model could incent smaller providers who are not yet participating in pay-for-performance, to be motivated for improvement.   Also, this model aligns with most other payment models now running in the community.

Recommendation 7

For the initial phase of implementation of SEGIP's incentive program, annual payment should be awarded as follows:

- If clinic site meets or exceeds the benchmark for the measure: $ 100 for each patient treated at that clinic (for the defined condition); **or**
- If clinic site attains the improvement target: $ 50 for each patient treated at that clinic (for the defined condition)

The per patient payments should be determined annually according to SEGIP budget requirements and consistency with other self-insured payers.

Rationale:

Studies show that even relatively modest rewards initially may be effective if providers know with certainty that the scope of the pay for performance effort, in terms of number of patients and payers involved, will increase in a relatively brief time. These levels would be consistent with the range that other payers are using in the community.

Comments:

This section of the report was generally not open for discussion in support of Anti-trust laws.

Recommendation 8

A point system to weight physician practice scores across all performance measures should be evaluated for implementation in future years.   Weighting of performance measures will be more important as additional measures are included in the system. This system should be based on population goals and concerns of the payers.

Rationale:

An aggregated scoring system will allow fine tuning of the weighting process in addition to setting targets or benchmarks for specific measures of care. The weighting allows the menu of measures to expand and the threshold specifications to remain stable while the points awarded for achieving targets for specific indicators are adjusted to prioritize the incentives allocation across all measures. The assumption of the equivalency of any gain or attainment reflected in the common practice of giving the same weight to each indicator may be acceptable in the short term development of quality incentives; but such an assumption does not reflect the true variation in value across quality indicators in terms of health gain, and this true variation should be addressed in a maturing and more comprehensive quality incentive system.

Comments:

Many found a point weighting system too complex to be communicated effectively. The hospital panel indicated that the current Hospital Compare Measure composite methodology may be too limiting, and felt the composite measures could be improved upon in terms of how best to incent high quality care and improvement. If implemented in future years, any weighting system should be determined and communicated before the measurement reporting process begins.

Recommendation 9

In the initial phase, address case mix and risk using the current inclusion and exclusion criteria, and by adjusting by major product type.

Inclusion and exclusion criteria established for each measure will be the first component of risk adjustment. In addition, performance scores will be adjusted by major product type (Public Programs, Medicare, Uninsured, and Commercial), using the following methodology for clinics:

> If a clinic site has: 1) a patient base with more than 20% of patients for that measure uninsured or in Public Programs, **and** 2) at least a representative sample of commercial members (n=20); that site's case mix will be adjusted against the calculated statewide average case mix for each measure.

Rationale:

(Note: This recommendation was included in the preliminary recommendations of the University of Minnesota and endorsed by the Payment Incentive Workgroup, but represenst a conflict of interest for MN Community Measurement, which would be the entity performing the risk adjustment.)

The issue for risk adjustment is how to be fair to those being assessed, without inadvertently establishing a policy that accepts lower quality care for populations with barriers. For the most part, one would expect that the treatment would be carried out irrespective of patient characteristics. For the measures that have been recommended for use by SEGIP for payment incentives, the results for the public programs and uninsured are on average lower across all providers. Therefore, a provider with a high number of these patients may not be recognized even though their result by population is the same as other groups that meet the benchmark. This adjustment method would address the case mix difference while still reflecting the actual results patients have received with that provider. Since the impact of case mix differences are small compared to the total variation across providers for these measures, small differences in case mix between providers will have little impact on the result. Therefore it is appropriate to adjust the result only for those providers with a significant public program and uninsured population. Also, if other payers choose to not implement this case mix adjustment, this method will allow SEGIP to align with them on the recognition for the majority of providers, and still have the potential to recognize a few additional safety net providers. This methodology is also feasible to implement for SEGIP by 2010.

 Comments:
One payer suggested that risk adjustment would mask the current transparency of pay for performance and jeopardize equal treatment to underserved populations. They commented that the solution rests in fixing the underlying payment system, not in the bonus payment system. Other payers were in favor of risk adjustment as it addresses the problem of poor performing patients affecting providers' scores, and felt using commercial scores alone was not appropriate.

Recommendation 10
The State should evaluate other risk adjustment methods for the new outcomes measures that may be included in its pay-for-performance program in the future. We recommend hospital risk adjustment continue to focus on co-morbidity and severity.

Rationale:

For Ambulatory measures we do not currently have reliable data nor well tested risk models for introducing severity as a risk adjuster in many care settings. In addition, we do not have data on some critical social factors that can serve as barriers to providing care or improving outcomes. This will become an increasingly important methodological issue as pay-for-performance increases in scale and scope and as performance improves for the less difficult to treat. This objective should be a key element of the second phase of pay-for-performance implementation.

Comments:
There was general agreement by all parties.

Recommendation 11

If feasible, the State should consider a formal Program Evaluation of this initiative after sufficient experience from the program has developed.

Rationale:
Formal Program Evaluation is valuable for an initiative of this scale. Results may help redefine aspects of the program, and potentially redirect valuable resources.

Comments:
One payer voiced concern over unintended consequences of this initiative, citing mixed historical research on pay for performance programs in general, the potential for rewarding negative behavior, and the possible limitation of innovation in the industry.

Next Steps

Minnesota Community Measurement will continue advising the MDH during the rulemaking process, which will begin following the review of this report. MNCM will proceed with implementation of the Quality Reporting and Payment System by educating physician clinics and hospitals about quality reporting requirements and the incentive payment system. Minnesota Community Measurement will begin collecting, validating, and summarizing data for public reporting, starting January 1, 2010.

Appendix A: Final Recommendations Report

<u>Quality Measures Inventory</u>

<u>Report to Minnesota Community Measurement</u>

Submitted by the

University of Minnesota

In collaboration with

Stratis Health

Minnesota Medical Association

Minnesota Hospital Association

December 5, 2008

Table of Contents

Section I – Introduction

The purpose of this report is to provide a comprehensive inventory of quality measures being developed and used in the U.S. for hospitals and clinics. The inventory will be a resource to support a prioritization process and decision making to determine new measures for hospitals and clinics that will be implemented by the State. The inventory is the result of combining and building on the previous related work of the following organizations:

- Agency for Health Care Quality and Research (AHRQ) National Quality Measures Clearinghouse (NQMC)
- Minnesota Community Measurement
- Stratis Health
- Minnesota Medical Association
- Minnesota Hospital Association

The University of Minnesota served to coordinate input from the sources above and compile the inventory. After we began to explore alternative approaches to developing the inventory, we determined that the AHRQ National Quality Measures Clearinghouse (NQMC) was a single resource that with its comprehensiveness, depth of information on specific measures, and ease of conducting searches for focused comparison of measures, had achieved most of the objectives of the State regarding an inventory, with the exception of a relatively small number of measures identified that were not included in the AHRQ data base and also information about measures in current use in Minnesota.

The NQMC includes over 1,400 developed measures and tracks over 500 measures under development. Each measure has been given a unique name and numerical ID. The information on the measures is extensive and addresses virtually all of the information requested by the State. NQMC even met our objective for an easily searchable data base for selecting measures by applying a large number of attributes of the measures. It includes measures in use by a number of entities such as the Institute for Clinical Improvement. NQMC includes an extensive collection of specialist quality measures that have been developed by professional societies and vetted through the AMA collaborative process. It includes measures on hospitals, nursing homes, and home health. It identifies the IOM aim addressed by each measure.

Extensive and detailed information that is provided for each measure captured all of the following attributes that was requested in the RFP, except for Minnesota use information.

Examples of attributes included in the RFP and addressed in the NQMC:

- Unique measure ID and name
- Institute of Medicine aim being addressed
- Name of measure in other measure sets
- Source /Initiative (e.g. NCQA)
- Clinical condition
- Population (age/gender/program/etc.)
- Part of delivery system being measured
- Description of the measure / Relationship to desired health outcome/ Evidence-base

- Domain: (Structure / Process / Outcome / Patient experience/ etc.)
- Data source(s)
- Current use status in US (examples of users and uses / Under development)
- Method for calculating the measure / inclusion and exclusion criteria
- Availability of state or national data that can be used to benchmark

Our conclusion is that since the inventory objectives are largely and competently addressed by NQMC, our tasks for the inventory should be to:

1) Present the measures and functionality of the NQMC and become skilled in using the NQMC data base to support the process of measure evaluation and prioritization
2) Identify measures that may be of interest, but are not in the NQMC
3) Identify measures in use in Minnesota
4) Build upon the NQMC by reviewing the feasibility and relevance of the measures

To support the entire University of Minnesota activity related to the MNCM project, we have retained a research assistant who is a master level student in health care administration with health care work experience. He has studied the extensive capabilities of the NQMC data base and will become fully trained in its use. We propose that this RA will be able to 1) train others in the use of NQMC and 2) conduct customized and timely searches of the NQMC to support the continuing work of the State through spring of 2009. The RA will be available to MNCM.

Inventory Report Organization

The inventory is organized as follows:

Section I – Introduction

Section II – Provides a brief introduction to quality measurement organizations nationally

Section III – Describes the National Quality Measures Clearinghouse (NQMC)

Section IV – Identifies measures not found in the NQMC

Section V – Identifies measures being used in Minnesota

Section VI – Presents high level issues to consider in evaluating measures for Minnesota

Appendix A – Provides examples of searches and measures comparisons using the National quality Measures Clearinghouse

Appendix B – Provides sample measures from the MMIC

Attachment I – Provides a lengthy, but the most abbreviated list available of the NMQM measures

Section II – Background

Quality performance measurement has been an organized effort of a number of national and regional organizations representing health care purchasers, consumers, policy decision makers, and providers. In recent years, the activities of these numerous entities have become more collaborative and coordinated. Now quality performance measures are generally reviewed by high level collaboratives of these organizations, regardless of where the measure originated historically. The desire to standardize measures crosses stakeholders. Providers need to have more coherence of measures. They need the measures that have passed thorough scrutiny. Purchasers and policymakers in our pluralistic health care purchasing system need to increase penetration within a provider's patient population in order to reach "critical mass" for change and to produce value-based incentive systems.

The following briefly describes selected key national organizations involved with health care quality measure development and evaluation:

1. *National Committee for Quality Assurance (NCQA)* – The National Committee for Quality Assurance (NCQA) is a private, not-for-profit organization dedicated to improving health care quality. NCQA has helped to build consensus around important health care quality issues by working with large employers, policymakers, doctors, patients and health plans to decide what's important, how to measure it, and how to promote improvement. NCQA develops quality standards and performance measures for a broad range of health care entities. It is with these measures and standards that organizations and individuals can identify opportunities for improvement.
Source: http://www.ncqa.org/tabid/675/Default.aspx

2. *National Quality Forum (NQF)* – The National Quality Forum (NQF) is a not-for-profit membership organization created to develop and implement a national strategy for health care quality measurement and reporting. Leaders in the public and private sector were prompted to create NQF as a mechanism to bring about national change in health care quality on patient outcomes, workforce productivity, and health care costs. The organizational members of the NQF work together to promote a common approach to measuring health care quality and fostering system-wide capacity for quality improvement.
Source: http://www.qualityforum.org/about/

3. *Agency for Healthcare Research and Quality (AHRQ)* – As a health services research arm of the U.S. Department of Health and Human Resources, the Agency for Healthcare Research and Quality (AHRQ) works with public and private sectors to build a knowledge base for what works, and for what does not work, in health and health care and to translate this knowledge into everyday practice and policymaking. AHRQ's main goals are to support improvements in health outcomes, develop strategies to strengthen quality measurement and improvement, and to identify healthcare strategies to improve health cost access, foster appropriate use, and reduce unnecessary expenditures in healthcare.
Source: http://www.ahrq.gov/about/whatis.htm

4. *Hospital Quality Alliance (HQA)* – In December 2002, the organizations representing America's hospitals joined with consumer representatives, physician and nursing organizations, employers and payers, oversight organizations and government agencies to launch the Hospital Quality Alliance (HQA). The HQA is a national public-private collaboration that is committed to making meaningful, relevant, and easily understood information about hospital performance accessible to the public. It also informs and encourages efforts to improve quality by using clinical quality, patient experience, equity, efficiency, and pricing information to spur positive changes in health care delivery.
Source: http://www.hospitalqualityalliance.org

5. *Ambulatory Quality Alliance (AQA)* – In 2004, the American Academy of Family Physicians (AAFP), the American College of Physicians (ACP), America's Health Insurance Plans (AHIP), and the Agency for Healthcare Research and Quality (AHRQ), joined together to determine how to most effectively and efficiently improve performance measurement, data aggregation, and reporting in the ambulatory care setting. Originally known as the Ambulatory Care Quality Alliance, the coalition is now known as the AQA alliance, because its mission has broadened to incorporate all areas of physician practice. AQA's mission and goals focus on areas that can help identify quality gaps, control skyrocketing cost trends, reduce confusion over redundant measures and alleviate administrative burdens in the marketplace.
Source: http://www.aqaalliance.org/default.htm

6. *The Joint Commission* is an independent, not-for-profit organization that accredits and certifies more than 15,000 health care organizations and programs in the United States. Joint Commission accreditation and certification is recognized nationwide as a symbol of quality that reflects an organization's commitment to meeting certain performance standards. The Joint Commission's mission is to continuously improve the safety and quality of care provided to the public through the provision of health care accreditation and related services that support performance improvement in health care organizations.
Source: http://www.jointcommission.org/AboutUs/

7. *American Medical Association (AMA)* – As the nation's largest association of physicians and medical students in the United States, the American Medical Association (AMA) advocates on the issues vital to the nation's health. AMA's mission is to promote the art and science of medicine and the betterment of public health. Their goal is to unite physicians nationwide to work on the most important professional and public health issues. By 2010, the AMA's goal will be to combine national Medicare and private health plan claims data and then use the data for public reporting of physician performance on quality and cost measures.
Source: http://www.ama-assn.org/ama/pub/category/1815.html

The national structure for quality measurement development, testing, and approval is centered on the National Quality Forum (NQF). NQF is a collaborative with support across the spectrum of stakeholders. Minnesota will need to consider the national priority agenda as it selects measures for the State. (Note: The measures being addressed by this collaborative are all included in the AHRQ NQMC data base.)

Included in the "National Priorities Partners" agenda setting process are the following organizations:

- National Partnership for Women and Families
- Consumers Union
- AARP
- AFL-CIO
- National Business Group on Health
- The Leapfrog Group
- Pacific Business Group on Health
- Chamber of Commerce
- Ambulatory Quality Alliance (AQA)
- Hospital Quality Alliance (HQA)
- Quality Alliance Steering Committee
- Alliance for Pediatric Quality
- AMA's Physician Consortium for Performance Improvement (PCPI)
- American Nurses Association
- American Board of Medical Specialties
- National Association of Community Health Centers
- Joint Commission
- National Committee for Quality Assurance
- Certification Commission for Healthcare Information Technology
- Centers for Disease Control and Prevention
- Centers for Medicare and Medicaid Services (CMS)
- Agency for Healthcare Research and Quality (AHRQ)
- National Institutes of Health
- National Governors Association
- America's Health Insurance Plans
- Institute for Healthcare Improvement
- Institute of Medicine (IOM)

The NQF national priority objectives are to: 1) center on high-leverage areas to achieve high return on investment, 2) harmonize efforts of "multiple groups" around common goals for improvement, and 3) emphasize the urgent need to drive fundamental change in delivery system.

NQF has recently determined the national priority areas for quality measurement for improvement:

1) *Patient and family engagement* – Engage patients and their families in managing health and making decisions about care
   Areas of focus:
   - Patient experience of care
   - Patient self-management

- Informed decision making

2) *Population health* – Improve the health of the population
   Areas of focus:
   - Healthy lifestyle behaviors
   - Preventive care
   - Community index to assess health status

3) *Safety* – Improve the safety and reliability of America's health care system
   Areas of focus:
   - Healthcare-associated infections
   - Serious adverse events
   - Mortality

4) *Care coordination* – Ensure patients receive well-coordinated care across all providers, settings, and levels of care
   Areas of focus:
   - Medication reconciliation
   - Preventable hospital readmissions
   - Preventable emergency department visits

5) *Palliative care* – Ensure patients receive well-coordinated care across all providers, settings, and levels of care
   Areas of focus:
   - Medication reconciliation
   - Preventable hospital readmissions
   - Preventable emergency department visits

6) *Overuse* – Eliminate overuse while ensuring the delivery of appropriate care
   Areas of focus:
   - Inappropriate medication use
   - Unnecessary lab tests
   - Unwarranted maternity care interventions
   - Unwarranted diagnostic procedures
   - Unwarranted procedures
   - Unnecessary consultations
   - Preventable emergency department visits and hospitalizations
   - Inappropriate non-palliative services at end of life
   - Potentially harmful preventive services with no benefit

(Source: NQF)

Distinct differences exist between hospital measures and reporting, and outpatient/clinic measures and reporting. The measurement of hospital quality performance has been largely driven by national requirements, especially by CMS and the Joint Commission, and a national system of data collection. To

be as cost-effective as possible, work on hospital measures will be built upon the national efforts. The measurement of clinic quality and performance has developed primarily from the local or state-based realm, and has only more recently had national efforts, such as CMS' PQRI program (Physician Quality Reporting Initiative). We will draw upon expertise and experience specific to Minnesota to bridge these two very different measurements and reporting realms.

Section III – Using the National Quality Measurement Clearinghouse Data Base

The AHRQ National Quality Measures clearinghouse is a remarkable resource for planners wishing to implement quality measures. The website describes the NQMC as follows:

> "The National Quality Measures Clearinghouse™ (NQMC), sponsored by the Agency for Healthcare Research and Quality (AHRQ) U.S. Department of Health and Human Services, is a database and Website for information on specific evidence-based health care quality measures and measure sets. NQMC is sponsored by AHRQ to promote widespread access to quality measures by the health care community and other interested individuals.
>
> The NQMC mission is to provide practitioners, health care providers, health plans, integrated delivery systems, purchasers and others an accessible mechanism for obtaining detailed information on quality measures, and to further their dissemination, implementation, and use in order to inform health care decisions. NQMC builds on AHRQ's previous initiatives in quality measurement, including the Computerized Needs-Oriented Quality Measurement Evaluation System (CONQUEST), the Expansion of Quality of Care Measures (Q-SPAN) project, the Quality Measurement Network (QMNet) project, and the Performance Measures Inventory"

Some limitations are that it does not include ALL measures we are interested in, including rural measures and others that MMA identified and the University determined are unique. The clearinghouse does not include application and relevance to Minnesota providers.

The following is a table which provides a description of the information included in the NQMC for each measure.

Complete Summary of NQMC Measure Attributes

Title	Identifies the title of the measure.	
Source(s)	Identifies the complete bibliographic source(s) for the measure as disseminated by the measure submitter(s).	
Primary Measure Domain	Classifies the major focus of the measure by one of the domains of care.	Choose one: • Access • Outcome • Patient Experience • Population Health • Process

Title	Identifies the title of the measure.	
		• Structure • Use of Services
Secondary Measure Domain	Identifies the secondary focus of the measure by domain of care (if applicable). *Does not apply to Use of Services and Population Health measures.*	Choose all that apply: • Access • Outcome • Patient Experience • Process • Structure
<td colspan="3" align="center">Brief Abstract</td>		
Description	Provides a concise statement of the specific aspects of health care, the patient population, providers, setting(s) of care, and time period that the measure addresses.	
Rationale	Identifies the rationale that briefly explains the importance of the measure (i.e., why it is used).	
Primary Clinical Component	Identifies the clinical aspect to which the measure refers, such as a structural feature, a clinical condition, a clinical process, a health outcome, and/or a patient characteristic. A combination of components may be identified (e.g., colorectal cancer; screening).	
Denominator Description	Provides the *general* specifications of any clinical component that is the basis for inclusions and exclusions in the denominator.	
Numerator Description	Provides the *general* specifications of any clinical component that is the basis for inclusions and exclusions in the numerator.	
<td colspan="3" align="center">Evidence Supporting the Measure</td>		
Evidence Supporting the Criterion of Quality	Describes the type(s) of supporting evidence appropriate for the measure domain. • For access measures, evidence that an association exists between the result of the access measure and the outcomes of, or satisfaction with, care. • For outcome measures, evidence that the outcome measure has been used to detect the impact of one or more clinical interventions. • For patient experience measures, evidence that an association exists between the measure of patient experience of health care and the values	Choose all that apply: • Unspecified • A clinical practice guideline or other peer-reviewed synthesis of the clinical evidence • A formal consensus procedure, involving experts in relevant clinical, methodological, and organizational sciences • A systematic review of the clinical literature (e.g., Cochrane Review) • Focus groups • One or more research studies published in a National Library of Medicine (NLM) indexed, peer-reviewed journal

Title	Identifies the title of the measure.	
	and preferences of individuals/the public. • For process measures, evidence that the measured clinical process has led to improved health outcomes. • For structure measures, evidence that an association exists between the structure measure and one of the four other domains of quality (e.g., access, outcome, patient experience, and process). Type of evidence includes published peer-reviewed studies, systematic reviews, and clinical practice guidelines, formal consensus procedures involving experts in relevant clinical, methodological, and organizational sciences. For patient experience measures, evidence should include focus groups involving patients and/or cognitive testing of the measure by patients. For access and structure measures, the consensus panel should also include other relevant stakeholders. *Does not apply to Population Health and Use of Services measures.*	
Evidence Supporting the Value of Monitoring the Aspect of Population Health	Describes the supporting evidence, if provided, for Population Health measures. *Does not apply to Access, Outcome, Patient Experience, Process, Structure or Use of Services measures.*	Choose all that apply: • No evidence is provided • A clinical practice guideline or other peer-reviewed synthesis of the clinical evidence • A formal consensus procedure, involving experts in relevant clinical, methodological, and organizational sciences • A systematic review of the clinical literature (e.g., Cochrane Review) • One or more research studies

Title	Identifies the title of the measure.	
		published in a National Library of Medicine (NLM) indexed, peer-reviewed journal
Evidence Supporting the Value of Monitoring Use of Service	Describes the supporting evidence if provided for Use of Services measures. *Does not apply to Access, Outcome, Patient Experience, Population Health, Process, or Structure measures.*	Choose all that apply: • No evidence is provided • A clinical practice guideline or other peer-reviewed synthesis of the clinical evidence • A formal consensus procedure, involving experts in relevant clinical, methodological, and organizational sciences • A systematic review of the clinical literature (e.g., Cochrane Review) • One or more research studies published in a National Library of Medicine (NLM) indexed, peer-reviewed journal
National Guideline Clearinghouse Link	Identifies link(s) to guideline summary(s) in the National Guideline Clearinghouse™ (NGC) where the measure was developed from an evidence-based guideline.	
Need for the Measure	Describes the type(s) of evidence that supports the need for the measure (i.e., why this measure was selected by the submitter).	Choose all that apply: • Unspecified • Overall poor quality for the performance measured • Use of this measure to improve performance • Variation in quality for the performance measured
	For Structure measures	Choose all that apply: • Unspecified • Overall insufficient capacity • Use of this measure to increase capacity • Variation in capacity
	For Use of Services measures	Choose all that apply: • Unspecified • Monitoring and planning • Variation in use of service
	For Population Health measures	Choose all that apply:

Title	Identifies the title of the measure.	
		• Unspecified • Monitoring health state(s) • Variation in health state(s)
Evidence Supporting Need for the Measure	Identifies references that support the assertions made regarding the need for the measure.	
State of Use	Identifies the status of the measure regarding its use within the past three years by health care organizations. Measure use can encompass current routine use, pilot testing, or still in use by organizations/entities although discontinued by the measure developer.	Choose one: • Unspecified • Current routine use • Pilot testing • Used, but developer discontinued
Current Use	Classifies the current use(s) of the measure by quality initiative and constituency (e.g., Internal quality improvement, Decision-making by consumers about health plan/provider choice). *The values "Internal quality improvement," "Collaborative inter-organizational quality improvement," "Quality of care research," "Decision-making by businesses about health-plan purchasing," "Decision-making by consumers about health plan/provider choice," and "Decision-making by health plans about provider contracting" may not be selected for Use of Services and Population Health measures.*	Choose all that apply: • Unspecified • Accreditation • Collaborative inter-organizational quality improvement • Decision-making by businesses about health-plan purchasing • Decision-making by consumers about health plan/provider choice • Decision-making by health plans about provider contracting • Decision-making by managers about resource allocation • External oversight/Department of Defense/TRICARE • External oversight/Indian Health Service • External oversight/Maternal and Child Health Bureau • External oversight/Medicaid • External oversight/Medicare • External oversight/Prison health care systems • External oversight/Regional, county, or city agencies

Title	Identifies the title of the measure.	
		• External oversight/State government program • External oversight/Veterans Health Administration • Federal health policymaking • Internal quality improvement • Monitoring and planning • Monitoring health state(s) • National reporting • Pay-for-performance • Quality of care research • State health policymaking
Application of the Measure in its Current Use		
Care Setting	Classifies the settings for which the measure applies.	Choose all that apply: • Unspecified • Ambulatory Care • Ancillary Services • Behavioral Health Care • Community Health Care • Emergency Medical Services • Home Care • Hospices • Hospitals • Long-term Care Facilities • Managed Care Plans • Physician Group Practices/Clinics • Rehabilitation Centers • Residential Care Facilities • Rural Health Care • Substance Use Treatment Programs/Centers
Professionals Responsible for Health Care	Classifies the professional(s) who is/are responsible for health care. *For all area health indicators, the value "Public Health Professionals" must be selected.*	Choose all that apply: • Unspecified • Advanced Practice Nurses • Allied Health Personnel • Chiropractors • Clinical Laboratory Personnel • Dentists • Dietitians

Title	Identifies the title of the measure.	
		• Emergency Medical Technicians/Paramedics • Measure is not provider specific • Nurses • Occupational Therapists • Pharmacists • Physical Therapists • Physician Assistants • Physicians • Podiatrists • Psychologists/Non-physician behavioral Health Clinicians • Public Health Professionals • Respiratory Care Practitioners • Social Workers • Speech-language Pathologists
Lowest Level of Health Care Delivery Addressed	Classifies the most discrete level of health care delivery to which the measure (in its current use) applies.	Choose one: • Unspecified • National • Regional • States • Counties or Cities • Metropolitan Statistical Areas/Health Services Areas • Multisite Health Care Organizations • Single Health Care Delivery Organizations • Group Clinical Practices • Individual Clinicians
Target Population Age	Describes the age range for the population measured. *Does not apply to Structure measures.*	
Target Population Gender	Classifies the target population by gender. *Does not apply to Structure measures.*	
Stratification by Vulnerable Populations	Describes the populations vulnerable to health care quality problems that are separately identified for sampling (e.g., Children, Homeless, Medically Uninsured). *Does not apply to Structure measures.*	

Title	Identifies the title of the measure.	
Incidence/Prevalence	Describes the occurrence in a population of the disease or condition or the structural feature associated with the primary clinical component.	
Evidence for Incidence/Prevalence	Identifies references documenting information provided in the Incidence/Prevalence field.	
Association with Vulnerable Populations	Describes the association of the primary clinical component within a population vulnerable to health care quality problems.	
Evidence for Association with Vulnerable Populations	Identifies references documenting information provided in the Association with Vulnerable Populations field.	
Burden of Illness	Describes the time course and amount of disability associated with the primary clinical component.	
Evidence for Burden of Illness	Identifies references documenting information provided in the Burden of Illness field.	
Utilization	Describes the utilization of resources due to the primary clinical component that may include hospital days, admissions/discharges, ambulatory care visits, tests, and procedures.	
Evidence for Utilization	Identifies references documenting information provided in the Utilization field.	
Costs	Describes the costs associated with the primary clinical component that may include per diem costs, or the cost of ambulatory care visits, tests, and procedures. In cases where costs for these items are not known, but charges are used as a proxy for cost.	
Evidence for Costs	Identifies references documenting information provided in the Costs field.	
Institute of Medicine National Health Care Quality Report Categories		
IOM Care Need	Classifies the measure into one of four Institute of Medicine (IOM) care need classifications where applicable. *Structure and Use of Services measures will always have the value "Not within an IOM Care Need."*	Choose all that apply to the primary clinical component: • Unspecified • Not within an IOM Care Need • End of Life Care • Getting Better • Living with Illness • Staying Healthy
IOM Domain	Classifies the measure into one or more of the Institute of Medicine (IOM) care domains where applicable. *Structure and Use of Services measures will always have the value "Not within an IOM Domain."* *The IOM Domain "Efficiency" can only be selected in conjunction with one of the other IOM Domains.*	Choose all that apply: • Unspecified • Not within an IOM Domain • Effectiveness • Efficiency • Equity • Patient-centeredness • Safety • Timeliness

Title	Identifies the title of the measure.	
Case Finding	Characterizes patients eligible for inclusion in the measure as users and/or nonusers of care. *Does not apply to Structure measures.*	Choose one: • Unspecified • Does not apply to this measure • Both users and nonusers of care • Users of care only
Description of Case Finding	Describes the procedure for determining whether a case is potentially eligible for inclusion in the denominator of a measure. Case finding establishes a sampling frame from which a more highly specified selection of cases will be made. *Does not apply to Structure measures.*	
Denominator Sampling Frame	Classifies the cases *potentially* eligible for inclusion in the denominator, from which a more highly specified selection of cases will be made. *Does not apply to Structure measures.*	Choose one: • Unspecified • Does not apply to this measure • Enrollees or beneficiaries • Geographically defined • Organizationally defined • Patients associated with provider
Denominator Inclusions/Exclusions	Describes the specific inclusion and exclusion criteria used to refine the denominator.	
Relationship of Denominator to Numerator	Designates whether all cases in the denominator are equally eligible to appear in the numerator. *Does not apply to Structure measures.* *Only the value "All cases in the denominator are not equally eligible to appear in the numerator" can be selected for Population Health measures.* *The value "Unspecified" cannot be selected for Access, Outcome, Patient Experience, Process, or Structure measures.*	Choose one: • Unspecified • Does not apply to this measure • All cases in the denominator are equally eligible to appear in the numerator • All cases in the denominator are not equally eligible to appear in the numerator
Denominator (Index) Event	Identifies the event or state that defines a patient as eligible for inclusion in the denominator.	Choose all that apply: • Unspecified • Does not apply to this measure • Clinical Condition

Title	Identifies the title of the measure.	
	Does not apply to Structure measures.	- Diagnostic Evaluation - Encounter - Institutionalization - Patient Characteristic - Provider Characteristic - Therapeutic Intervention
Denominator Time Window	Classifies the time period (in association with the denominator [index] event) in which patients are reviewed for inclusion in the denominator. *Does not apply to Structure measures.*	Choose one: - Unspecified - Does not apply to this measure - Time window brackets index event - Time window follows index event - Time window is a fixed period of time - Time window is a single point in time - Time window precedes index event
Numerator Inclusions/Exclusions	Describes the specific inclusion and exclusion criteria used to refine the numerator. This field will be used to further describe the metric (if necessary).	
Measure Results Under Control of Health Care Professionals, Organizations and/or Policymakers	Designates whether measure results are somewhat or substantially under the control of the health care professionals, organizations and policymakers to whom the measure applies. *The value "Unspecified" cannot be selected for Access, Outcome, Patient Experience, Process, or Structure measures.*	Choose one: - Unspecified - The measure results are somewhat or substantially under the control of the health care professionals, organizations and/or policymakers to whom the measure applies. - The measure results are not under the control of the health care professionals, organizations and/or policymakers to whom the measure applies.
Numerator Time Window	Identifies the time period in which patients are reviewed for inclusion in the numerator.	Choose one: - Unspecified - Does not apply to this measure - Encounter or point in time

Title	Identifies the title of the measure.	
		• Episode of care • Fixed time period • Institutionalization
Data Source	Identifies the data source(s) necessary to implement the measure.	Choose all that apply: • Unspecified • Administrative and laboratory data • Administrative and medical records data • Administrative and pharmacy data • Administrative and provider data • Administrative data • Administrative data and clinician survey • Administrative data and patient survey • Clinician survey • Laboratory data • Medical record • National public health data • Patient survey • Pharmacy data • Provider data • Registry data • Special or unique data • State public health data
	For Structure measures, the following are possible data sources:	<> • Administrative data • Clinician survey • National public health data • Provider data • Special or unique data • State public health data
Level of Determination of Quality	Identifies the level at which quality can be assessed, i.e., at the individual patient level or the aggregate patient level. *Does not apply to Structure, Population*	Choose one: • Does not apply to this measure • Individual Case • Not Individual Case

Title	Identifies the title of the measure.	
	Health, and Use of Services measures.	
Outcome Type	Classifies the type of outcome for Outcome measures. *Applies only when "Outcome" is selected as a Primary or Secondary Measure Domain.*	Choose one: • Unspecified • Does not apply to this measure • Adverse Outcome • Clinical Outcome • Functional Status • Health Risk State or Behavior • Proxy for Outcome • Quality of Life Measure
Type of Health State	Classifies the type of health state for Population Health measures. *Applies only to Population Health measures.*	Choose one: • Unspecified • Does not apply to this measure • Adverse Health State • Functional Status • Health Risk State or Behavior • Health State not otherwise specified
Pre-existing Instrument Used	Identifies all pre-existing instruments, such as a standardized survey instrument, used in implementing the measure.	
Computation of the Measure		
Scoring	Identifies the method used to score the measure.	Choose one: • Unspecified • Categorical Variable • Continuous Variable • Count • Frequency Distribution • Non-weighted Score/Composite/Scale • Rate • Ratio • Weighted Score/Composite/Scale
Interpretation of Score	Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score. Applies to Process, Outcome, Access, Experience, and Structure measures.	Choose one: • Better quality is associated with a higher score • Better quality is associated with a lower score • Better quality is associated with a score falling within a defined interval

Title	Identifies the title of the measure.	
		• Passing score defines better quality
	Applies to Population Health measures.	Choose one: • Unspecified • A higher score is desirable • A lower score is desirable • A score falling within a defined interval is desirable
	Applies to all Use of Services measures.	• Undetermined
Allowance for Patient Factors	Identifies the type of analytic considerations made for the measure based on patient factors or characteristics. *Does not apply to Structure measures.*	Choose all that apply: • Unspecified • Does not apply to this measure • Analysis by high-risk subgroup (stratification on vulnerable populations) • Analysis by subgroup (stratification on patient factors, geographic factors, etc.) • Case-mix adjustment • Paired data at patient level • Risk adjustment devised specifically for this measure/condition • Risk adjustment method widely or commercially available
Description of Allowance for Patient Factors	Describes the analytic considerations made for the measure based on the patient factors and characteristics. *This field will not display if either "Unspecified" or "Does not apply to this measure" is selected in the "Allowance for Patient Factors" field.*	
Standard of Comparison	Classifies the type and time frame of the comparison according to whether the comparison is external (at a given point-in-time or of a time trend), internal or to a prescriptive standard. *The specific nature of the "prescriptive standard" (e.g., "pass/fail") will be described in the corresponding text field.*	Choose all that apply: • Unspecified • Does not apply to this measure • External comparison at a point in time • External comparison of time trends • Internal time comparison • Prescriptive standard

Title	Identifies the title of the measure.
Prescriptive Standard	Describes the prescriptive standard(s) used for comparison of measure results.
Evidence for Prescriptive Standard	Identifies references documenting information provided in the Prescriptive Standard field. *Does not apply to Use of Services and Population Health measures.*
Evaluation of Measure Properties	
Extent of Measure Testing	Describes the extent of testing of the measure including reliability and/or validity testing.
Evidence for Reliability/Validity Testing	Identifies references documenting reliability/validity testing as described in the Extent of Measure Testing field.
Original Title	Identifies the original name of the measure as stated in the original measure documentation.
Measure Collection	Identifies the name of the collection of measures to which the measure belongs (if applicable).
Measure Set Name	Identifies the name of the measure set to which the measure belongs (if applicable).
Measure Subset Name	Identifies the name of the subset to which the measure belongs (if applicable).
Composite Measure Name	Identifies the name of the composite measure to which the measure belongs (if applicable).
Submitter	Identifies the organization(s) that submitted the measure to NQMC.
Developer	Identifies the organization(s) that developed the measure.
Funding Source(s)	Identifies source(s) of funding to the organization(s) for developing the measure(s) or measure set(s)/collection(s).
Composition of the Group that Developed the Measure	Describes the composition of the group/committee that developed the measure(s) or measure set(s)/collection(s), including professional degrees and affiliations, and lists the names of individual committee members, where given.
Financial Disclosures/Other Potential Conflicts of interest	Records and makes publically available disclosed relationships between individuals of the measure development committee/group/individual and companies or organizations that could potentially influence that individual's contribution to the development of the measure(s) or measure set(s)/collection(s).
Endorser	Identifies the organization(s) that have endorsed the measure.
Included In	Identifies the inclusion of a measure in specified measure initiatives (e.g., National Healthcare Quality Report (NHQR), National Healthcare Disparities Report (NHDR), Home Health Compare).
Adaptation	Identifies that the measure has been adapted *from* another measure(s).
Parent Measure	Identifies the name(s) of all the measures from which the current measure was adapted. The name of each "parent" measure's developer follows in parentheses.
Release Date	Identifies the date that the measure was first released by the submitting organization (this could be the date first issued or published).
Revision Date	Identifies the date of the most recent revision to the measure and/or the documentation by the submitting organization (if applicable).

Title	Identifies the title of the measure.
Measure Status	Identifies whether the measure is the current release or an update.
Source(s)	Identifies the complete bibliographic source(s) for the measure as disseminated by the measure submitter(s).
Measure Availability	Identifies contact information for requesting the measure documentation. Where possible, information regarding electronic (including hypertext links to the full-text) and print copies is provided.
Companion Documents	Identifies companion documents that are relevant to the measure. These companion documents are not necessarily available within NQMC.
NQMC Status	Identifies when the measure was completed or revised by ECRI, and verified by the submitting organization(s).
Copyright Statement	Provides the copyright statement of the organization that submitted the measure.
Disclaimer	

(Source: AHRQ)

For a complete listing of NQMC measures, see Attachment I. Because of the large number of measures, the list is quite long; however, it is a minimum data list including only the NQMC name of the measure and the initiative associated with the measure. The list can be scanned to find types of measures of interest. While connected to the internet and with an electronic version of the report, one can "Control click" on any of the measures to link to a "complete summary" of the measure. An example of a complete summary is provided below.

Example: Complete Summary of a Measure

TITLE

Health plan members' experiences: percentage of adult health plan members who reported how often their personal doctor communicated well.

SOURCE(S)

- CAHPS® health plan survey and reporting kit 2007. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2006 Nov 14. Various p.
- CAHPS®: Surveys and tools to advance patient-centered care [https://www.cahps.ahrq.gov/default.asp]. [internet]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); [updated 2008 Feb 21]; [accessed 2007 Sep 05]. [3 p].

Measure Domain

PRIMARY MEASURE DOMAIN

Patient Experience

The validity of measures depends on how they are built. By examining the key building blocks of a measure, you can assess its validity for your purpose. For more information, visit the Measure Validity page.

SECONDARY MEASURE DOMAIN

Does not apply to this measure

Brief Abstract

DESCRIPTION

This measure is used to assess the percentage of respondents who indicated how often ("Never," "Sometimes," "Usually," or "Always") their personal doctor:

- explained things in a way that was easy for them to understand.
- listened carefully to them.
- showed respect for what they had to say.
- spent enough time with them.

The "How Well Doctors Communicate" composite measure is based on four questions in the CAHPS Health Plan Survey 4.0 (Adult Questionnaire).

Note: A composite score is calculated in which a higher score indicates better quality. Composite scores are intended for consumer-level reporting. Additionally, frequency distributions are available for plans or providers to use for quality improvement purposes.

RATIONALE

The Agency for Healthcare Research and Quality (AHRQ) (then called the Agency for Health Care Policy and Research, or AHCPR) initiated the CAHPS program in October 1995 to develop standardized survey tools for obtaining and reporting information on consumers' experiences with health care. The CAHPS consortium began by developing the CAHPS Health Plan Survey, an integrated set of carefully tested and standardized questionnaires and report formats that can be used to produce meaningful, reliable, and comparable information about the experiences of consumers enrolled in health plans. The CAHPS Health Plan Survey is designed to generate information that consumers can use to choose health plans, that purchasers can use to assess the value of services they buy, and that health plans can use to assess their performance and improve their products and services. As AHRQ had intended, the survey can be used with all types of health insurance consumers--including Medicaid recipients, Medicare beneficiaries, and those who are commercially insured-- and across the full range of health care delivery systems, from fee-for-service to managed care plans. The instruments also capture information about special groups, including individuals with chronic conditions and disabilities and families with children.

The National Committee for Quality Assurance (NCQA) requires health plans to submit measures from the CAHPS Health Plan Survey as part of their HEDIS submission and for accreditation purposes.

PRIMARY CLINICAL COMPONENT

Health care; members' experiences; physician communication

DENOMINATOR DESCRIPTION

Health plan members age 18 years and older who answered the "How Well Doctors Communicate" questions on the CAHPS Health Plan Survey 4.0 (Adult Questionnaire) (see the "Description of Case Finding" and the "Denominator Inclusions/Exclusions" fields in the Complete Summary)

NUMERATOR DESCRIPTION

The number of "Never," "Sometimes," "Usually," or "Always" responses on the "How Well Doctors Communicate" questions (see the related "Numerator Inclusions/Exclusions" field in the Complete Summary)

Evidence Supporting the Measure

EVIDENCE SUPPORTING THE CRITERION OF QUALITY

A formal consensus procedure involving experts in relevant clinical, methodological, and organizational sciences
One or more research studies published in a National Library of Medicine (NLM) indexed, peer-reviewed journal

Evidence Supporting Need for the Measure

NEED FOR THE MEASURE

- Use of this measure to improve performance
- Variation in quality for the performance measured

EVIDENCE SUPPORTING NEED FOR THE MEASURE

- What consumers say about the quality of their health plans and medical care: The National CAHPS Benchmarking Database. 2007 CAHPS health plan survey chartbook. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2007 Dec. 41 p.

State of Use of the Measure

STATE OF USE

Current routine use

CURRENT USE

Accreditation

Decision-making by businesses about health-plan purchasing

Decision-making by consumers about health plan/provider choice

External oversight/Department of Defense/TRICARE

External oversight/Medicaid

External oversight/Medicare

External oversight/State government program

Internal quality improvement

National reporting

Quality of care research

Application of Measure in its Current Use

CARE SETTING

Managed Care Plans

PROFESSIONALS RESPONSIBLE FOR HEALTH CARE

Physicians

LOWEST LEVEL OF HEALTH CARE DELIVERY ADDRESSED

Single Health Care Delivery Organizations

TARGET POPULATION AGE

Age greater than or equal to 18 years

TARGET POPULATION GENDER

Either male or female

STRATIFICATION BY VULNERABLE POPULATIONS

Unspecified

Characteristics of the Primary Clinical Component

INCIDENCE/PREVALENCE

Unspecified

ASSOCIATION WITH VULNERABLE POPULATIONS

Unspecified

BURDEN OF ILLNESS

Unspecified

UTILIZATION

Unspecified

COSTS

Unspecified

Institute of Medicine National Healthcare Quality Report Categories

IOM CARE NEED

End of Life Care

Getting Better

Living with Illness

Staying Healthy

IOM DOMAIN

Patient-centeredness

Data Collection for the Measure

CASE FINDING

Both users and nonusers of care

DESCRIPTION OF CASE FINDING

Health plan members age 18 years and older, who have been enrolled in:

- the commercial plan for 12 months or longer, with no more than one 45-day break in enrollment during the 12 months

OR

- a Medicaid plan or product for 6 months or longer, with no more than one 30-day break in enrollment during the 6 months.

DENOMINATOR SAMPLING FRAME

Enrollees or beneficiaries

DENOMINATOR INCLUSIONS/EXCLUSIONS

Inclusions

Health plan members age 18 years and older who answered the "How Well Doctors Communicate" questions on the CAHPS Health Plan Survey 4.0 (Adult Questionnaire). Include refusals, non-response, and bad addresses/phone numbers.

Exclusions

- Individuals with coverage other than primary health coverage, such as a dental-only plan
- Deceased
- Ineligible (not enrolled in the plan)

RELATIONSHIP OF DENOMINATOR TO NUMERATOR

All cases in the denominator are equally eligible to appear in the numerator

DENOMINATOR (INDEX) EVENT

Patient Characteristic

DENOMINATOR TIME WINDOW

Time window precedes index event

NUMERATOR INCLUSIONS/EXCLUSIONS

Inclusions

The number of "Never," "Sometimes," "Usually," or "Always" responses on the "How Well Doctors Communicate" questions

From the responses, a composite score is calculated in which a higher score indicates better quality.

Note: Include all completed questionnaires. A questionnaire is considered complete if responses are available for 10 or more of a selected list of key CAHPS items. Refer to the original measure documentation for more information.

Exclusions

Unspecified

MEASURE RESULTS UNDER CONTROL OF HEALTH CARE PROFESSIONALS, ORGANIZATIONS AND/OR POLICYMAKERS

The measure results are somewhat or substantially under the control of the health care professionals, organizations and/or policymakers to whom the measure applies.

NUMERATOR TIME WINDOW

Fixed time period

DATA SOURCE

Administrative data

Patient survey

LEVEL OF DETERMINATION OF QUALITY

Not Individual Case

PRE-EXISTING INSTRUMENT USED

Unspecified

Computation of the Measure

SCORING

Non-weighted Score/Composite/Scale

INTERPRETATION OF SCORE

Better quality is associated with a higher score

ALLOWANCE FOR PATIENT FACTORS

Analysis by subgroup (stratification on patient factors, geographic factors, etc.)

Case-mix adjustment

DESCRIPTION OF ALLOWANCE FOR PATIENT FACTORS

CAHPS recommends adjusting the data for respondent age, education, and general health status.

If the sample size is sufficient, responses may be analyzed for specific sub-populations, such as respondents with chronic conditions.

STANDARD OF COMPARISON

External comparison at a point in time

External comparison of time trends

Internal time comparison

Evaluation of Measure Properties

EXTENT OF MEASURE TESTING

The CAHPS Health Plan Survey has probably been tested more completely than any previously used consumer survey. There are two different and complementary approaches to assessing the reliability and validity of a questionnaire (1) cognitive testing, which bases its assessments on feedback from interviews with people who are asked to react to the survey questions, and (2) psychometric testing, which bases its assessments on the analysis of data collected by using the questionnaire. Although many existing consumer questionnaires about health care have been tested primarily or exclusively using a psychometric approach, the CAHPS team views the combination of cognitive and psychometric approaches as essential to producing the best possible survey instruments. Consequently, both methods have been included in the development of the CAHPS survey.

The cognitive testing method provided useful information on respondents' perceptions of the response task, how respondents recalled and reported events, and how they interpreted specified reference periods. It also helped identify words that could be used to describe health care providers accurately and consistently across a range of consumers (e.g., commercially insured, Medicaid, fee-for-service, managed care, lower socioeconomic status [SES], middle SES, low literacy, higher literacy) and helped explore whether key words and concepts included in the core questions worked equally well in both English and Spanish.

The CAHPS consortium also tested each CAHPS reporting composite in focus groups with plan members. Cognitive interviews with consumers were conducted to ensure that the reporting composites and their labels were easily understood. Psychometric analyses using data collected during pilot tests were also conducted. These analyses indicated that both the composites and the items in each composite were reliable and valid measures of members' experiences. In addition, items in each reporting composite were tested and found to be internally consistent. For example, reliability coefficients (Cronbach's alpha) in one pilot test involving four health plans using the instrument that most resembled the

final CAHPS 2.0 instrument ranged from a low of 0.68 for the "Getting Needed Care" composite to a high of 0.90 for the "How Well Doctors Communicate" composite. These composites are positively associated with members' ratings of overall care provided by doctors and nurses and ratings of health plans.

In addition, the CAHPS development team, together with researchers from the National Committee on Quality Assurance (NCQA), conducted a detailed comparative analysis of the items in the CAHPS questionnaire and NCQA's Member Satisfaction Survey (MSS) from the fall of 1997 to the spring of 1998. These questionnaires were merged to form the 2.0 version of the CAHPS questionnaire. This testing is noteworthy because it was so extensive and because of the wide array of techniques used. These included focus groups, in-depth cognitive testing, pilot studies, methodological experiments, and large demonstration studies, such as the demonstrations in Washington State, Kansas, and New Jersey. NCQA also worked with the CAHPS consortium to conduct field tests of the 4.0 instrument with six health plans in Spring 2005.

EVIDENCE FOR RELIABILITY/VALIDITY TESTING

- CAHPS®: Surveys and tools to advance patient-centered care [https://www.cahps.ahrq.gov/default.asp]. [internet]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); [updated 2008 Feb 21]; [accessed 2007 Sep 05]. [3 p].

Identifying Information

ORIGINAL TITLE

How well doctors communicate.

MEASURE COLLECTION

CAHPS Health Plan Survey

MEASURE SET NAME

CAHPS Health Plan Survey 4.0, Adult Questionnaire

SUBMITTER

Agency for Healthcare Research and Quality

DEVELOPER

Agency for Healthcare Research and Quality

CAHPS Consortium

Centers for Medicare & Medicaid Services

FUNDING SOURCE(S)

Unspecified

COMPOSITION OF THE GROUP THAT DEVELOPED THE MEASURE

Unspecified

FINANCIAL DISCLOSURES/OTHER POTENTIAL CONFLICTS OF INTEREST

Unspecified

ENDORSER

National Quality Forum

ADAPTATION

Measure was not adapted from another source.

RELEASE DATE

1997 Mar

REVISION DATE

2006 Nov

MEASURE STATUS

This is the current release of the measure.

This measure updates a previous version: CAHPS® Health Plan Survey and Reporting Kit 2002 (3.0 Version). Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2002. This previous version of the CAHPS Health Plan Survey remains available for use through 2007. Therefore, NQMC will retain the CAHPS Health Plan Survey 3.0 version on the NQMC Web site.

SOURCE(S)

- CAHPS® health plan survey and reporting kit 2007. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2006 Nov 14. Various p.
- CAHPS®: Surveys and tools to advance patient-centered care [https://www.cahps.ahrq.gov/default.asp]. [internet]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); [updated 2008 Feb 21]; [accessed 2007 Sep 05]. [3 p].

MEASURE AVAILABILITY

The individual measure, "How Well Doctors Communicate," is published in the "CAHPS Health Plan Survey and Reporting Kit 2007." This Kit may be downloaded at the CAHPS Survey Users Network Web site. See the related QualityTools summary.

COMPANION DOCUMENTS

The following are available:

- What consumers say about the quality of their health plans and medical care: The National CAHPS Benchmarking Database. 2007 CAHPS health plan survey chartbook. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2007 Dec. 41 p. This document is available in Portable Document Format (PDF) from the CAHPS Web site. See the related QualityTools summary.
- CAHPS user resources: project implementation resources. [Web site]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); [updated 2005 Dec 29]; [accessed 2007 Jan 25]. Available from the CAHPS Web site.
- CAHPS community: the report card compendium. [Web site]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); [updated 2006 Nov 14]; [accessed 2007 Jan 25]. Available from the CAHPS Web site. CAHPS survey and reporting kits. Project profiles. [Web site]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); [updated 2007 Feb 5]; [accessed 2007 Jan 25]. Available from the CAHPS Web site.
- CAHPS survey and reporting kits. Reporting resources: downloadable documents. [Web site]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); [updated 2007 Jun 12]; [accessed 2007 Jan 25]. Available from the CAHPS Web site.
- Edgman-Levitan S, Shaller D, McInnes K, Joyce R, Coltin KL, Cleary PD, Rybowski L, editor(s). The CAHPS improvement guide. Practical strategies for improving the patient care experience. Cambridge (MA): Department of Health Care Policy, Harvard Medical School; 2003 Oct 1. 156 p. This document is available in (PDF) from the CAHPS Web site.

NQMC STATUS

This NQMC summary was completed by ECRI on April 24, 2007. The information was verified by the measure developer on June 15, 2007.

COPYRIGHT STATEMENT

No copyright restrictions apply.

(Source: AHRQ)

Conducting Measure Searches and Measure Comparisons Using the NQMC Data Base

NQMC will be updated continuously and accepts submission of new measures from many sources. There are numerous ways to search and select on measures for comparison. A tutorial is offered to assist users in techniques for searching the stat base.

See Appendix A for three examples of searches:
1) Diabetes nephropathy measures
2) Diabetes measures for 13-18 year olds
3) Cardiologist measures

Section IV: Measures of Interest Not Found in the NQMC

We cross-referenced measures reported from a large number of state and national sources to identify any unique measures that were not among the vast number included in the NQMC. We found a relatively small set. In some cases, the measure was listed as unique if it broke down the NQMC related measure into more detailed sub-categories. The example of this reported below is the set measures related to antibiotic use for surgery. Stratis Health-reported measures define the measure for each of a number specific types of surgery; whereas NQMC includes the same measure but primarily aggregates across types of surgery. Because some Minnesota hospitals use the surgery type-specific measures, they are considered unique and relevant.

For other measures, namely the rural hospital measures and a set of detailed functional status measures for home health care reported below, the measures themselves are unique and, again, relevant for Minnesota. These measures are not repeated in the next section where we report on additional measures in use in Minnesota whether unique or included in the NQMC.

Unique Measures (Source: Stratis Health)

Rural Hospital Measures

Prescribing Practices	Read back verbal orders
	Admission orders reconciled with home meds
	Review of order by pharmacist within 24 hours
	Pharmacist rounds with physicians
Documenting Practices	Handwritten MAR
	Electronic MAR from pharmacy software
	MAR verified against order before drug prep
Medication Acquisition Practices	RN/LPN responsible for obtaining new medications (M-F Day)
	RN/LPN responsible for obtaining new medications (Weekend Day)
	RN/LPN responsible for obtaining new medications (M-F Night)
	Independent double check in pharmacy (M-F Day)
	Majority of oral medications in unit dose form
	Automated dispensing cabinet in use
Administering Practices	Meds routinely selected/administered by same person
	Two identifiers (excluding room no.) used to establish patient identity
	Unopened unit dose verified with MAR at bed
Medication Error Reporting Practices	Error reports NOT placed in personnel files
	NCC MERP taxonomy used to categorize error severity
	Near misses routinely reported
	Medication errors discussed at medication safety committee
	Conducted root cause analysis within the last year
Safe Culture Practices	Aggregate medication error data compared to external database
	Aggregate medication error data shared with hospitals of similar size
	Survey of patient safety culture conducted in the past year
	Harmful errors disclosed to patients/families
	Accredited by JCAHO

Pharmacist Support	Pharmacist employed by hospital	
	Pharmacist onsite 5 or fewer hours per week	
	Contract with local community pharmacist	
	Current pharmacy vacancy	
	Report they lack patient volume to support full time pharmacist	
	Limited financial resources	
	Shortage of pharmacists	
	Stakeholders disagree concerning need for pharmacy report	
Emergency Department Chest Pain/AMI Assessment Measures	Time to ECG	Time of arrival at ED until time of first 12 lead ECG. Includes pre-hospital ECG at 0 minutes. ACC and A HA standard of 10 minutes is used for the standard
	Aspirin within 24 hours	Proportion of CP/AMI patients in the ED without aspirin contraindications who received aspirin within 24 hours before or after hospital arrival
	Time to Thrombolytics	Proportion of ED AMI patients with ST elevation on ECG whose time from hospital arrival to thrombolysis is 30 minutes or les
Emergency Department Trauma Vital Signs Measure	Emergency Department Trauma Vital Signs	Proportion of trauma patients with systolic blood pressure, pulse rate, or respiratory rate documented on arrival to the ED and at least hourly (or until ER patient is released, admitted or transferred).
Emergency Department transfer time and communication	ED transfer time and communication	Number of information elements sent with transfer patients in 7 categories (pre-transfer communication, patient identification, vital signs, medication - related information, physician generated information, nurse generated information, and procedures and tests).
Emergency Department transfer time and communication	Patient Time in ED (longer than 2 hours)	
	Patient Time in ED (longer than 4 hours)	
	Condition Categories:	
	Brain injured	
	Burns	
	Crushing Injury	
	Foreign body	
	Fracture	
	Internal injury	
	Open wounds	
	Arrived by Ambulance	
	Discharge Status	
	Admitted to this hospital	
	Discharged to home	
	Discharged to ICF	
Emergency Department Transfer Communication Measure	Left AMA	
	ED transfer communication	
	Transferred to short term general hospital	
	Administrative information:	
	1. Nurse communication with receiving hospital staff	
	2. Physician communication with receiving professional staff	
	Patient information	
	1. Name	
	2. Address	

    3. Age

    4. Gender

    5. Contact information for significant others

    6. Insurance information

Vital Signs

    1. Pulse

    2. Respiration

    3. Blood Pressure

    4. Temperature

    5. Oxygen level

    6. Glasgow score

    7. Apgar score

Medication communication

    1. Medication history

    2. Medications given (MAR)

    3. Allergies

Physician documentation

    1. Physician's history and physical

    2. Physician's orders and reason for transfer

Nurse documentation

    1. Nurse documentation: interventions/response to care

    2. Impairments

    3. Immobility

    4. Respiratory support given

    5. Oral restriction

    6. Catheters

Tests and procedures

    1. Tests and procedures done

    2. Tests and procedures sent

Home Health Measures

OBQI Outcome
Measures

    Outcome Measures

    Improvement in grooming

    Stabilization in grooming

    Improvement in dressing upper body

    Improvement in dressing lower body

    Improvement in bathing

    Stabilization in bathing

    Improvement in toileting

    Improvement in transferring

Stabilization in transferring

Improvement in ambulation

Improvement in eating

Improvement in light meal preparation

Stabilization in light meal preparation

Improvement in laundry

Stabilization in laundry

Improvement in housekeeping

Stabilization in housekeeping

Improvement in shopping

Stabilization in shopping

Improvement in phone use

Stabilization in phone use

Improvement in management of oral meds

Stabilization in management of oral meds

Improvement in speech and language

Stabilization in speech and language

Improvement in pain interfering w/ activity

Improvement in number of surgical wounds

Improvement in status surgical wounds

Improvement in dyspnea

Improvement in urinary tract infection

Improvement in urinary incontinence

Improvement in bowel incontinence

Improvement in cognitive functioning

Stabilization in cognitive functioning

Improvement confusion frequency

Improvement in anxiety level

Stabilization in anxiety level

Improvement in behavioral problem frequency

Any emergent care

Utilization Measures

Discharge to the Community

Acute care hospitalization

Additional Unique Measures (Source: Stratis Health)

Domain	Measure	Description
Outpatient (AMI) and Chest Pain	OP-5	Median Time to ECG
AMI	AMI-3	ACEI or ARB for LVSD
AMI	AMI-8	Median Time to Primary PCI
AMI	AMI-8a	Primary PCI Received Within 90 Minutes of Hospital Arrival
AMI	AMI-9	Inpatient Mortality

AMI	AMI-T1a	LDL-Cholesterol Assessment (Optional Test Measure)
AMI	AMI-T2	Lipid-Lowering Therapy at Discharge (Optional Test Measure)
HF	HF-2	Evaluation of LVS Function
PN	PN-6a	Initial Antibiotic Selection for CAP in Immunocompetent – ICU Patient
PN	PN-6b	Initial Antibiotic Selection for CAP Immunocompetent – Non ICU Patient
SCIP	SCIP-Inf-1a	Prophylactic Antibiotic Received Within One Hour Prior to Surgical Incision - Overall Rate
SCIP	SCIP-Inf-1b	CABG
SCIP	SCIP-Inf-1c	Other Cardiac Surgery
SCIP	SCIP-Inf-1d	Hip Arthroplasty
SCIP	SCIP-Inf-1e	Knee Arthroplasty
SCIP	SCIP-Inf-1f	Colon Surgery
SCIP	SCIP-Inf-1g	Hysterectomy
SCIP	SCIP-Inf-1h	Vascular Surgery
SCIP	SCIP-Inf-2a	Prophylactic Antibiotic Selection for Surgical Patients - Overall Rate
SCIP	SCIP-Inf-2b	CABG
SCIP	SCIP-Inf-2c	Other Cardiac Surgery
SCIP	SCIP-Inf-2d	Hip Arthroplasty
SCIP	SCIP-Inf-2e	Knee Arthroplasty
SCIP	SCIP-Inf-2f	Colon Surgery
SCIP	SCIP-Inf-2g	Hysterectomy
SCIP	SCIP-Inf-2h	Vascular Surgery
SCIP	SCIP-Inf-7	Colorectal Surgery Patients with Immediate Postoperative Normothermia
SCIP	SCIP-Card-2	Surgery Patients on Beta-Blocker Prior to Arrival /Received a Beta-Blocker Perioperative
SCIP	SCIP-VTE-1	Surgery Patients with Recommended Venous Throboembolism Prophylaxis Ordered
SCIP	SCIP-VTE-2	Surgery Patients Who Received Appropriate Venous Thromboembolism Prophylaxis within 24 Hours Prior to Surgery to 24 Hours After Surgery
PR	PR-1	VBAC
PR	PR-3	Third or Fourth Degree Laceration
CAC	CAC-1b	Relievers for Inpatient Asthma (age 2 through 4 years)
CAC	CAC-1c	Relievers for Inpatient Asthma (age 5 through 12 years)
CAC	CAC-1d	Relievers for Inpatient Asthma (age 13 through 17 years)
CAC	CAC-2a	Systemic Corticosteroids for Inpatient Asthma (age 2 through 17 years) – Overall Rate
CAC	CAC-2b	Systemic Corticosteroids for Inpatient Asthma (age 2 through 4 years)
CAC	CAC-2c	Systemic Corticosteroids for Inpatient Asthma (age 5 through 12 years)
CAC	CAC-2d	Systemic Corticosteroids for Inpatient Asthma (age 13 through 17 years)
CAC	CAC-3	Home Management Plan of Care (HMPC) Document Given to Patient/Caregiver

Section V – Measures in Use in Minnesota

Minnesota has long been a pioneer in the development and implementation of quality measures for quality improvement and public reporting. Health plans were using claims data to assess physician practice breast cancer screening rates as early as 1988. These efforts preceded NCQA and other national efforts. In fact, Minnesota health plans have been at the forefront of quality measurement activities

nationally. In addition, Minnesota health plans and purchasers have collaborated to develop standards for assessing and reporting. The most recent and successful of these is the Minnesota Community Measurement organization that has become a national leader in collaborative measurement of physician quality performance and is extending its work beyond both health plans and claims data as part of the RWJ Aligning Forces for Quality program and related direct data submissions from physician practices. MNCM also supports the implementation of the Bridges to Excellence program in the State.

Another collaborative that has produced measures to support the implementation of evidence-based guidelines in the region is the Institute for Clinical Systems Improvement (ICSI). ICSI has developed a number of measures associated with monitoring adherence with ICSI guidelines. These measures have influenced national efforts and have been used to adapt national measures for regional use.

Stratis Health is also an innovative quality performance and management organization that serves as the QIO for the region. Stratis Health and the Minnesota Hospital Association came together in 2005 to launch the Minnesota Hospital Quality Report. This web-based report card includes measures of clinical care in heart failure, pneumonia, heart attack, and surgical care, including the innovative Appropriate Care Measure, a patient-focused measure that provides a way of looking at whether a patient received ALL of the "appropriate" or "right care" (recommended treatments) that they should have received, based on their clinical condition. The report recently added a measure of patients' experiences in the hospital, as assessed by HCAHPS.

Stratis Health completed in 2005 the Rural Measures Special Study for the Federal Centers for Medicare & Medicaid Services, through which Stratis Health led a multi-state field test of new rural-relevant hospital emergency department measures, and facilitated a national technical Expert Panel, resulting in recommendations to CMS that were enacted in the 2008 Outpatient Prospective Payment System.

In addition, health plans have adapted quality measures or developed new measures for their pay-for-performance programs. With this history of quality measurement and collaboration, Minnesota is well positioned to continue the advancement of quality performance measurement of physician practices and hospitals over the next few years to achieve a community-wide, all-payer standard for performance assessment and reporting.

Currently, nursing homes and home health agencies are reporting quality performance information for the Medicare Home Health and Nursing Home Compare programs. Additional quality assessment of nursing homes and hospitals is also being conducted. Hospitals in the state are likewise reporting quality performance information for the Medicare Hospital Compare program. Critical-access hospitals (CAHs) and other hospitals previously not permitted to report their outpatient quality information will soon be able to do so. In addition, Minnesota Hospital Association and the Minnesota Department of Health are sponsoring an adverse event reporting initiative that also includes developing and adopting protocols for reducing adverse events. The four adverse events being reported for improvement include: wrong body part surgery, retained foreign objects, falls, and pressure ulcers. The Minnesota Hospital Association is also participating in an Agency for Healthcare Research and Quality sponsored two-year project that allows clinical lab data to be paired with administrative billing data.

MHA reports that, "Hospitals already submit billing data to MHA. Any hospital can be part of this new initiative by agreeing to also submit their clinical lab data. Once the new lab data and billing data are merged, a more sophisticated severity adjustment system can be applied, and hospital performance on quality and patient safety measures can be more accurately analyzed. The new merged data will also

help hospitals double-check their accuracy for coding conditions present on admission."

The longest standing and most developed public reporting has been of health plans initially and now physician practices - particularly in primary care. While physician practice level quality performance reporting has been well underway in Minnesota, Medicare has only recently implemented its voluntary Physician Quality Reporting Initiative (PQRI) which advances Medicare plans to implement public reporting and pay-for-performance for physician practices. Physician practices have been the primary component of the health care delivery system being assessed by MN Community Measurement, health plans, and purchasers. The Minnesota Medical Association has compiled an inventory of quality measures being used in Minnesota for pay-for performance and other measurement purpose. The following table is a summary table of measures that was updated in 2008. Additional MMA tables provide extensive detail about the measures.

Minnesota Medical Association Measure Summary
(*Indicates measure is included in the NQMC)

Measure	BCBS MN 2008 Recognizing	Bridges to Excellence 2007-2008	HealthPartners Partners in Progress	HealthPartners Partners in Excellence	Medica Performance-Based	Medica Choice Care Quality	Preferred One Updated 5/28/8	UCare P4P program Updated	DHS Reporting requirements	QCare	MN Community Measurement	2008 CMS PQRI Updated
1. Acute Bronchitis*												x
2. Advanced Care Plan*												x
3. Asthma	x						x				x	x
4. Avoidance of Antibiotic Treatment in Adults with Acute Bronchitis*							x				x	
5. Board Maintenance of Certification	x											
6. Body mass index (BMI) and weight management plan*	x		x	x								x
7. Cancer – Breast Cancer*												X
8. Cancer – Chemotherapy plan*												x
9. Cancer – Chemotherapy for colon cancer*												x
10. Cancer – Chronic Lymphocytic Leukemia*												x
11. Cancer – Multiple Myeloma*												x
12. Cancer - Myelodysplastic Syndrome (MDS) and Acute Leukemias*												x
13. Cancer – Prostate Cancer*												x
14. Cardiovascular - Acute Myocardial Infarction*										x		x
15. Cardiovascular – Non-traumatic Chest pain*												x
16. Cardiovascular -	x									x		x

Measure	BCBS MN 2008 Recognizing	Bridges to Excellence 2007-2008	HealthPartners in Progress	HealthPartners in Excellence	Medica Performance-Based	Medica Choice Care Quality	Preferred One Updated 5/28/8	UCare P4P program Updated	DHS Reporting requirements	QCare	MN Community Measurement	2008 CMS PQRI Updated
Congestive Heart Failure*												
17. Cardiovascular - Congestive heart failure program*			x	x								
18. Cardiovascular – Coronary Artery Bypass Graph*												x
19. Cardiovascular – Coronary Artery Disease (Optimal cardiac care)*	x	x		x	x		x	x	x	x	x	x
20. Cardiovascular - Recurrent atrial Fibrillation*	x											
21. Carpal tunnel release			x									
22. Child and teen check up								x				
23. Child developmental screening incentive*									x			
24. Child mental health screening incentive									x			
25. Chronic Kidney Disease*												x
26. Chronic Obstructive Pulmonary Disease (COPD)*												x
27. Community Acquired Bacterial Pneumonia*												x
28. Critical Care												x
29. Depression Care*	x	x		x			x				x	x
30. Depression Symptom Assessment tool*			x									
31. Diabetes*	x	x		x	x	x	x	x	x	x	x	x
32. Discectomy for acute disc herniation			x									
33. Ear Care – Acute Otitis Externa*	x											x
34. Ear Care – Otitis Media with Effusion*												x
35. Elder Health Evaluation									x			
36. Electronic Clinical Data Reporting*			x									
37. End Stage Renal Disease*												x
38. Eye Care – Primary Open Angle Glaucoma*												x
39. Eye Care – Age related Macular Degeneration*												x
40. Eye Care – Diabetic Retinopathy*												x
41. Functional status*				x								
42. Gastroesophageal Reflux Disease												x

Measure	BCBS MN 2008 Recognizing	Bridges to Excellence 2007-2008	HealthPartners Partners in Progress	HealthPartners Partners in Excellence	Medica Performance-Based	Medica Choice Care Quality	Preferred One Updated 5/28/8	UCare P4P program Updated	DHS Reporting requirements	QCare	MN Community Measurement	2008 CMS PQRI Updated
(GERD)*												
43. Generic drug use*	x		x	x								
44. Generic Prescribing Provider Decision Support			x									
45. Health information technology (HIT)*	x		x								To be added in 2008	x
46. Hepatitis C*												X*
47. Hypertension							x				x	
48. Immunization rate – adolescents*										x		
49. Immunization rate – Children*							x	x	x	x	x	
50. Immunizations-Adults*												x
51. Innovations in Health Care*				x								
52. Language and race documentation			x									
53. Low back pain–Pain management			x									
54. Medication Reconciliation*												x
55. Meniscectomy Arthroscopy			x									
56. MNCM Direct Data Submission Participation			x									
57. Osteoarthritis*												x
58. Osteoporosis*												x
59. Pain assessment*												x
60. Palliative Care Program*			x									
61. Pathology												x
62. Patient satisfaction/ experience*				x							x	
63. PeriOperative Care*												x
64. Pharyngitis*							x				x	x
65. Pneumonia*				Included in Preventive Services Composite Measure						x		
66. Rheumatoid Arthritis*												x
67. Safety Composite Assessment			x									
68. Screening - Blood lead level*			x			x		x	x		x	
69. Screening - Breast cancer*				x		x		x	x	x	x	x

58

Measure	BCBS MN 2008 Recognizing	Bridges to Excellence 2007-2008	HealthPartners Partners in Progress	HealthPartners Partners in Excellence	Medica Performance-Based	Medica Choice Care Quality	Preferred One Updated 5/28/8	UCare P4P program Updated	DHS Reporting requirements	QCare	MN Community Measurement	2008 CMS PQRI Updated
70. Screening - Cervical cancer*				x			x	x	x	x	x	
71. Screening – Chlamydia*	x			Included in Preventive Services Composite Measure	x	x	x	x	x	x	x	
72. Screening – Clinical Depression*												x
73. Screening – Cognitive Impairment*												x
74. Screening - Colorectal cancer*				Included in Preventive Services Composite Measure			x	x			x	x
75. Screening -Composite cancer	x			Included in Preventive Services Composite Measure			x				x	
76. Screening – Future Falls Risk*												x
77. Screening – Preventative Services Composite - Adults				x								
78. Screening – Preventative Services Composite - Pediatrics				x								
79. Screening - Standardized alcohol abuse screen*	x			Included in Preventive Services Composite Measure								
80. Screening and Intervention Process for ED			x									
81. Spinal Surgery	x											

Measure	BCBS MN 2008 Recognizing	Bridges to Excellence 2007-2008	HealthPartners Partners in Progress	HealthPartners Partners in Excellence	Medica Performance-Based	Medica Choice Care Quality	Preferred One Updated 5/28/8	UCare P4P program Updated	DHS Reporting requirements	QCare	MN Community Measurement	2008 CMS PQRI Updated
82. Stroke and Stroke Rehab*												x
83. Syncope*												x
84. Tobacco – Identification*			x									x
85. Tobacco - Assisting smokers to quit*	x											x
86. Total hip or total knee replacement	x										In development 2008	
87. Total joint antibiotic prophylaxis*	x											
88. Upper respiratory Infection -Appropriate treatment for children*							x				x	x
89. Urinary Incontinence*												x
90. Well child visits – Infants								x		x	Retired 2007	
91. Well child visits – 3-6 years old								x	x		Retired 2007	

Source: Minnesota Medical Association)

Minnesota Community Measurement Measures (NCQA HEDIS / ICSI)

Minnesota Community Measurement measures are identified in the table above and in the detailed measure descriptions found on the MMA detailed measures tables.

The following lists MNCM measures:

MNCM measures
- Asthma
- Cancer Screening:
  - Breast
  - Cervical
  - Colorectal
  - Cancer Screening Combined*(Ages 50–80)*
- Childhood Immunization
- Chlamydia Screening
- Controlling High Blood Pressure
- Optimal Diabetes Care*
- Pharyngitis
- Upper Respiratory Infection
- Optimal Vascular Care

- Optimal Coronary Artery Disease
- (CAD) Care**
- Depression**
- Patient Experience***
- Health Information Technology**

Measures recently retired:
- Childhood Immunization *(Combo 2)*
- Depression Medication Management
- High Blood Pressure – *(old target)*
- Optimal Diabetes Care *(old targets)*
- Well Child Visits

Measures being considered for the future:
- CAD re-named Optimal Vascular Care**
- Asthma all-or-none measure**
- Specialty measure
- Cost of care / resource use
- Other care settings

MNCM is also implementing new depression measures that include outcome as well as process measures:

1. <u>Percentage of Adult Population Diagnosed with Major Depression or Dysthymia:</u> Adults (ages 18 and older) with a diagnosis of depression. Depression is common, with a lifetime risk for major depressive disorder of 7%-12% for men and 20%-25% for women *(U.S. Department of Health and Human Services Public Health Service, 1993)*. The depression codes for these measures will be:

   - 296.2 - Major Depressive disorder, single episode
   - 296.3 – Major depressive disorder, recurrent episode
   - 300.4 – Dysthymic Disorder

   <u># adult patients with depression (296.2x, 296.3x and 300.4)</u>
   total # adult patients
     Typical population statistics in primary care:
     Major Depression (296.2 and 296.3) – 5-9% prevalence women, 2-3% prevalence for Dysthymia
     (300.4) = 3% point prevalence in population

2. <u>Percentage of Adult Population with Depression NOS Diagnosis</u>: Adults (ages 18 and older) with a **diagnosis of 311 – Depressive disorder**, not elsewhere classified. PHQ-9 Response and Remission rates have not been validated for this diagnosis, and thus this information will be collected only for determining the proportion of patients given this diagnosis. *These patients will not be included in the subsequent measures.*

   <u># of adult patients with diagnosis of depression not elsewhere classified (311)</u>

total adult patients
   Typical population statistics in primary care:
   Depression Disorder NOS, 311 = 11% of population

3. <u>Percentage of Adult Population who had PHQ-9 at Baseline:</u> Percent of patients with a diagnosis of depression (296.2, 296.3 or 300.4) with a completed PHQ–9 at first (index) contact (+ 30 days) where depression was coded in the measurement period. Contact is defined as an office visit, telephone call, or e-visit with any practitioner.

    <u># adult pts with depression (296.2x, 296.3x and 300.4) who had a PHQ-9 administered</u>
    # adult patients with depression (296.2x, 296.3x and 300.4)

4. <u>Number of Adult Patients with depression and a PHQ-9 > 9 at Index Contact:</u> The number of patients with a PHQ-9 of > 9 at the index contact. Index contact is defined as the starting visit associated with a contact date in which the patient has a PHQ-9 score > 9 and ICD9 codes identifying the patient as having major depression or dysthymia.) This number serves as the denominator for measuring patient improvement at six and twelve months

Depression Outcome Measures:

The following table is a list of measures that can be calculated based on the direct data submission that occurs.

*   For the 2009 BTE rewards program, the only measure that will be used is the six- month remission rate defined as a six-month PHQ-9 score of < 5.
*   Response is defined as a 50% or more reduction of PHQ-9 score
*   Remission is defined as a PHQ-9 score of less than 5
*   PHQ-9 scores will be included if they are plus or minus 30 days of the point of measurement. For example a patient's index contact date is 2/15/2008. The six month date from this time would be 8/15/2008, but the patients contact date for PHQ-9 is 8/27/2008; this PHQ-9 score and date would be accepted because it is within the 60 day grace period.

Measure	Calculation
PHQ-9 follow-up assessment at six months	<u># adult pts with depression & PHQ-9 > 9 who have 6 month PHQ-9 (+/- 30 days)</u> # adult pts with depression with index contact PHQ-9 > 9
PHQ-9 with a 50 % or more decrease in score (response) at six months	<u># adult pts with > = 50% decrease in PHQ-9 score at 6 months</u> # adult pts with depression with index contact PHQ-9 > 9
PHQ-9 score < 5 (remission) at six months Note: BTE Measure	<u># adult pts with a PHQ-9 score < 5 at 6 months</u> # adult pts with depression with index contact PHQ-9 > 9
PHQ-9 follow-up assessment at twelve months	<u># adult pts with depression & PHQ-9 > 9 who have 12 month PHQ-9 (+/- 30 days)</u> # adult pts with depression with index contact PHQ-9 > 9
PHQ-9 with a 50 % or more decrease in score (response) at twelve months	<u># adult pts with > = 50% decrease in PHQ-9 score at 12 months</u> # adult pts with depression with index contact PHQ-9 > 9

| PHQ-9 score < 5 (remission) at twelve months | # adult pts with a PHQ-9 score < 5 at 12 months |
| | # adult pts with depression with index contact PHQ-9 > 9 |

(Source: Minnesota Community Measurement)

Professional Liability Prevention Measures

Another potential source of quality measures is the liability prevention activities of physician malpractice insurers. This approach has also been considered by the Colorado Guideline Collaborative using measures from their physician-owned malpractice consortium in Colorado. In Minnesota, this opportunity is being forwarded by Midwest Medical Insurance Company (MMIC). As a physician-owned medical professional liability insurance company, they have the acceptance of the physicians and work in their interest to help practices and institutions avoid mistakes that can lead to lawsuits.

MMIC has developed criteria by which they assess the level of malpractice risk clinics and hospitals face due to their internal systems, policies and procedures, and have developed recommended systems to help minimize those risks. The criteria that help liability insurers determine whether a healthcare facility is at low or high risk for malpractice claims can also be used to help assess the quality and patient safety levels of the facility.

Attached in Appendix B is list of risk management criteria and possible measures of those criteria and is intended only as a sample of what could be developed. Also is a risk management self-assessment that includes many other criteria that could be considered.

Section VI – Prioritizing Measures for Minnesota – Overview

With such a large menu of new measures from which to select, the process of evaluating measures must be guided by the usual concerns with measure validity, reliability,"action-ability", attribution fairness, adequate risk adjustment, and population health impact. In addition, for Minnesota, the NQMC measures and measures of interest from other sources will need to be reviewed with consideration of the state's objectives and unique measurement capabilities.

National priorities and related national evaluations of measures often apply to Minnesota, but some national conclusions will be less applicable. We will need to recognize where national priorities and conventional wisdom regarding a measure's feasibility and acceptability are based on applying nationally focused, least common denominator assumptions about data limitations or the feasibility of achieving a community-wide coordinated action related to performance measurement.

The NQMC includes extensive information about the evidence-base and current uses of measures. Some of that information will require a Minnesota-specific review. An example would be for measures that have typically not been considered for use at the individual provider level because of small sample size for a specific provider through one payer. Were such a caveat raised nationally, Minnesota, with MNCM and its legislated statewide objective, may have fewer constraints. Another constraint on national priorities, including Medicare, is the limitations of claims/administrative data. The CMS PQRI reporting includes intermediate outcomes reported on a HCFA 1500 claims using G codes. This is a workable "work-around" to overcome the limitations of claims data; however, Minnesota may have greater opportunity to implement outcome measurement because physician practices can report clinical data directly. Similarly, the MHA clinical data project sponsored by AHRQ is another example of advanced data availability potential. It is important to note that the clinical data not only allows for more extensive outcomes measurement, but also allows for improved risk adjusters as well.

Appendices:

<u>Examples Search and Comparisons from NQMC</u>

1) Measures of Diabetic Nephropathies

Diseases (MeSH Category) - List all 1184 measures
Endocrine System Diseases - List all 122 measures
Diabetes Mellitus - List all 103 measures
Diabetes Complications - List all 26 measures
Diabetic Nephropathies - List all 8 measures

Display results 1 to 8 of 8

☑ Adult diabetes: percentage of patients who received any test for microalbuminuria. National Diabetes Quality Improvement Alliance 2003 May. NQMC:000606

☑ Adult diabetes: percentage of patients with at least one test for microalbumin during the measurement year; or who had evidence of medical attention for existing nephropathy (diagnosis of nephropathy or documentation of microalbuminuria or albuminuria). National Diabetes Quality Improvement Alliance 2003 May. NQMC:000600

☑ Adult diabetes: percentage of patients with no urinalysis or urinalysis with negative or trace urine protein, who received a test for microalbumin. National Diabetes Quality Improvement Alliance 2003 May. NQMC:000607

☑ Comprehensive diabetes care: percentage of members 18 through 75 years of age with diabetes mellitus (type 1 and type 2) who had a nephropathy screening test or evidence of nephropathy. National Committee for Quality Assurance 2007 Jul. NQMC:002775

☑ Diabetes mellitus: hospital admission rate for long-term complications. Agency for Healthcare Research and Quality 2007 Mar. NQMC:003105

☑ [Diabetes mellitus: percent of eligible patients with a diagnosis of diabetes mellitus having a nephropathy screening test during the past year or documented evidence of nephropathy.](#) Veterans Health Administration
2007 Oct. NQMC:003814

☑ [Diabetes mellitus: the percentage of patients with diabetes who have a record of estimated glomerular filtration rate (eGFR) or serum creatinine testing in the previous 15 months.](#)
British Medical Association
National Health System (NHS) Confederation
2006 Feb. NQMC:001920

☑ [Diabetes mellitus: the percentage of patients with diabetes who have a record of micro-albuminuria testing in the previous 15 months (exception reporting for patients with proteinuria).](#)
British Medical Association
National Health System (NHS) Confederation
2006 Feb. NQMC:001919

---

2) Comparison of diabetes measures based on age of population (ages 13-18 years old)

Measure Comparison

	Diabetes mellitus: the percentage of patients with diabetes whose last measured total cholesterol within the previous 15 months is 5 mmol/l or less.	Diabetes mellitus: the percentage of patients with diabetes in whom the last HbA1c is 7.5 or less (or equivalent test/reference range depending on local laboratory) in the previous 15 months.	Diabetes mellitus: the percentage of patients with diabetes in whom the last HbA1c is 10 or less (or equivalent test/reference range depending on local laboratory) in the previous 15 months.	Diabetes mellitus: the percentage of patients with diabetes in whom the last blood pressure is 145/85 or less.	Diabetes mellitus: the percentage of patients with diabetes whose notes record body mass index (BMI) in the previous 15 months.
Title					
Measure Collection	Quality and Outcomes Framework Indicators	Quality and Outcomes Framework Indicators	Quality and Outcomes Framework Indicators	Quality and Outcomes Framework Indicators	Quality and Outcomes Framework Indicators
Submitter	British Medical Association	British Medical Association	British Medical Association	British Medical Association	British Medical Association

	National Health System (NHS) Confederation	National Health System (NHS) Confederation	National Health System (NHS) Confederation	National Health System (NHS) Confederation	National Health System (NHS) Confederation
Developer	British Medical Association National Health System (NHS) Confederation	British Medical Association National Health System (NHS) Confederation	British Medical Association National Health System (NHS) Confederation	British Medical Association National Health System (NHS) Confederation	British Medical Association National Health System (NHS) Confederation
Funding Source(s)	The expert panel who developed the indicators are entirely funded by a grant from the English Department of Health.	The expert panel who developed the indicators are entirely funded by a grant from the English Department of Health.	The expert panel who developed the indicators are entirely funded by a grant from the English Department of Health.	The expert panel who developed the indicators are entirely funded by a grant from the English Department of Health.	The expert panel who developed the indicators are entirely funded by a grant from the English Department of Health.
Composition of the Group that Developed the Measure	The main indicator development group is based in the National Primary Care Research and Development Centre in the University of Manchester. They are: Professor Helen Lester NPCRDC, MB BCH MD, Dr. Stephen Campbell, NPCRDC, PhD, Dr. Umesh Chauhan, NPCRDC, MB BS, PhD. Others involved	The main indicator development group is based in the National Primary Care Research and Development Centre in the University of Manchester. They are: Professor Helen Lester NPCRDC, MB BCH MD, Dr. Stephen Campbell, NPCRDC, PhD, Dr. Umesh Chauhan, NPCRDC, MB BS, PhD. Others involved in the development of individual indicators are: Professor Richard Hobbs, Dr. Richard McManus, Professor Jonathan	The main indicator development group is based in the National Primary Care Research and Development Centre in the University of Manchester. They are: Professor Helen Lester NPCRDC, MB BCH MD, Dr. Stephen Campbell, NPCRDC, PhD, Dr. Umesh Chauhan, NPCRDC, MB BS, PhD. Others involved in the development of individual indicators are: Professor Richard Hobbs, Dr. Richard McManus, Professor Jonathan	The main indicator development group is based in the National Primary Care Research and Development Centre in the University of Manchester. They are: Professor Helen Lester NPCRDC, MB BCH MD, Dr. Stephen Campbell, NPCRDC, PhD, Dr. Umesh Chauhan, NPCRDC, MB BS, PhD. Others involved	The main indicator development group is based in the National Primary Care Research and Development Centre in the University of Manchester. They are: Professor Helen Lester NPCRDC, MB BCH MD, Dr. Stephen Campbell, NPCRDC, PhD, Dr. Umesh Chauhan, NPCRDC, MB BS, PhD. Others involved

	in the development of individual indicators are: Professor Richard Hobbs, Dr. Richard McManus, Professor Jonathan Mant, Dr. Graham Martin, Professor Richard Baker, Dr. Keri Thomas, Professor Tony Kendrick, Professor Brendan Delaney, Professor Simon De Lusignan, Dr. Jonathan Graffy, Dr. Henry Smithson, Professor Sue Wilson, Professor Claire Goodman, Dr. Terry O'Neill, Dr. Philippa Matthews, Dr. Simon Griffin, Professor Eileen Kaner.	Mant, Dr. Graham Martin, Professor Richard Baker, Dr. Keri Thomas, Professor Tony Kendrick, Professor Brendan Delaney, Professor Simon De Lusignan, Dr. Jonathan Graffy, Dr. Henry Smithson, Professor Sue Wilson, Professor Claire Goodman, Dr. Terry O'Neill, Dr. Philippa Matthews, Dr. Simon Griffin, Professor Eileen Kaner.	Mant, Dr. Graham Martin, Professor Richard Baker, Dr. Keri Thomas, Professor Tony Kendrick, Professor Brendan Delaney, Professor Simon De Lusignan, Dr. Jonathan Graffy, Dr. Henry Smithson, Professor Sue Wilson, Professor Claire Goodman, Dr. Terry O'Neill, Dr. Philippa Matthews, Dr. Simon Griffin, Professor Eileen Kaner.	in the development of individual indicators are: Professor Richard Hobbs, Dr. Richard McManus, Professor Jonathan Mant, Dr. Graham Martin, Professor Richard Baker, Dr. Keri Thomas, Professor Tony Kendrick, Professor Brendan Delaney, Professor Simon De Lusignan, Dr. Jonathan Graffy, Dr. Henry Smithson, Professor Sue Wilson, Professor Claire Goodman, Dr. Terry O'Neill, Dr. Philippa Matthews, Dr. Simon Griffin, Professor Eileen Kaner.	in the development of individual indicators are: Professor Richard Hobbs, Dr. Richard McManus, Professor Jonathan Mant, Dr. Graham Martin, Professor Richard Baker, Dr. Keri Thomas, Professor Tony Kendrick, Professor Brendan Delaney, Professor Simon De Lusignan, Dr. Jonathan Graffy, Dr. Henry Smithson, Professor Sue Wilson, Professor Claire Goodman, Dr. Terry O'Neill, Dr. Philippa Matthews, Dr. Simon Griffin, Professor Eileen Kaner.
Financial Disclosures/Other Potential Conflicts of Interest	None for the main indicator development group.	None for the main indicator development group.	None for the main indicator development group.	None for the main indicator development group.	None for the main indicator development group.
Release Date	2004 Apr	2004 Apr	2004 Apr	2004 Apr	2004 Apr
Revision Date	2006 Feb	2006 Feb	2006 Feb	2006 Feb	2006 Feb

Description	This measure is used to assess the percentage of patients with diabetes whose last measured total cholesterol within the previous 15 months is 5 mmol/l or less.	This measure is used to assess the percentage of patients with diabetes in whom the last HbA1c is 7.5 or less (or equivalent test/reference range depending on local laboratory) in the previous 15 months.	This measure is used to assess the percentage of patients with diabetes in whom the last HbA1c is 10 or less (or equivalent test/reference range depending on local laboratory) in the previous 15 months.	This measure is used to assess the percentage of patients with diabetes in whom the last blood pressure is 145/85 or less.	This measure is used to assess the percentage of patients with diabetes whose notes record body mass index (BMI) in the previous 15 months.
Rationale	Diabetes mellitus is one of the common endocrine diseases affecting all age groups with over one million people in the United Kingdom (UK) having the condition. Effective control and monitoring can reduce mortality and morbidity. Much of the management and monitoring of diabetic patients, particularly patients with Type 2 diabetes is undertaken by the general practitioner and members of the primary care team. This	Diabetes mellitus is one of the common endocrine diseases affecting all age groups with over one million people in the United Kingdom (UK) having the condition. Effective control and monitoring can reduce mortality and morbidity. Much of the management and monitoring of diabetic patients, particularly patients with Type 2 diabetes is undertaken by the general practitioner and members of the primary care team. This measure is one of sixteen Diabetes Mellitus measures. The Diabetes Mellitus indicators are based on widely	Diabetes mellitus is one of the common endocrine diseases affecting all age groups with over one million people in the United Kingdom (UK) having the condition. Effective control and monitoring can reduce mortality and morbidity. Much of the management and monitoring of diabetic patients, particularly patients with Type 2 diabetes is undertaken by the general practitioner and members of the primary care team. This measure is one of sixteen Diabetes Mellitus measures. The Diabetes Mellitus indicators are based on widely	Diabetes mellitus is one of the common endocrine diseases affecting all age groups with over one million people in the United Kingdom (UK) having the condition. Effective control and monitoring can reduce mortality and morbidity. Much of the management and monitoring of diabetic patients, particularly patients with Type 2 diabetes is undertaken by the general practitioner and members of the primary care team. This	Diabetes mellitus is one of the common endocrine diseases affecting all age groups with over one million people in the United Kingdom (UK) having the condition. Effective control and monitoring can reduce mortality and morbidity. Much of the management and monitoring of diabetic patients, particularly patients with Type 2 diabetes is undertaken by the general practitioner and members of the primary care team. This

measure is one of sixteen [Diabetes Mellitus](#) measures. The [Diabetes Mellitus](#) indicators are based on widely recognised approaches to the care of diabetes. Detailed guidelines for health professionals are published by Diabetes UK and by SIGN - the Scottish Intercollegiate Guidelines Network. The SIGN website contains detailed evidence tables, and links to published articles. The English National Service Framework for Diabetes also includes details of the evidence behind a range of recommendations. The National Institute for Health and Clinical Excellence (NICE) has also	recognised approaches to the care of diabetes. Detailed guidelines for health professionals are published by Diabetes UK and by SIGN - the Scottish Intercollegiate Guidelines Network. The SIGN website contains detailed evidence tables, and links to published articles. The English National Service Framework (NSF) for Diabetes also includes details of the evidence behind a range of recommendations. The National Institute for Health and Clinical Excellence (NICE) has also published guidance on a number of aspects of diabetic control. The indicators for diabetes are generally those which would be expected to be done, or checked in an annual review. There is no requirement on the general practitioner (GP) practice to carry out all these	recognised approaches to the care of diabetes. Detailed guidelines for health professionals are published by Diabetes UK and by SIGN - the Scottish Intercollegiate Guidelines Network. The SIGN website contains detailed evidence tables, and links to published articles. The English National Service Framework for Diabetes also includes details of the evidence behind a range of recommendations. The National Institute for Health and Clinical Excellence (NICE) has also published guidance on a number of aspects of diabetic control. The indicators for diabetes are generally those which would be expected to be done, or checked in an annual review. There is no requirement on the general practitioner (GP) practice to carry out all these	measure is one of sixteen [Diabetes Mellitus](#) measures. The [Diabetes Mellitus](#) indicators are based on widely recognised approaches to the care of diabetes. Detailed guidelines for health professionals are published by Diabetes UK and by SIGN - the Scottish Intercollegiate Guidelines Network. The SIGN website contains detailed evidence tables, and links to published articles. The English National Service Framework for Diabetes also includes details of the evidence behind a range of recommendations. The National Institute for Health and Clinical Excellence (NICE) has also	measure is one of sixteen [Diabetes Mellitus](#) measures. The [Diabetes Mellitus](#) indicators are based on widely recognised approaches to the care of diabetes. Detailed guidelines for health professionals are published by Diabetes UK and by SIGN - the Scottish Intercollegiate Guidelines Network. The SIGN website contains detailed evidence tables, and links to published articles. The English National Service Framework for Diabetes also includes details of the evidence behind a range of recommendations. The National Institute for Health and Clinical Excellence (NICE) has also

published guidance on a number of aspects of diabetic control. The indicators for diabetes are generally those which would be expected to be done, or checked in an annual review. There is no requirement on the general practitioner (GP) practice to carry out all these items (e.g., retinal screening), but it is the practice's responsibility to ensure that they have been done. This set of indicators relates to both Type 1 and Type 2 diabetes. Although the care of patients with Type 1 diabetes may be shared with specialists, the general practitioner would still be expected to ensure that appropriate annual checks	items (e.g., retinal screening), but it is the practice's responsibility to ensure that they have been done. This set of indicators relates to both Type 1 and Type 2 diabetes. Although the care of patients with Type 1 diabetes may be shared with specialists, the general practitioner would still be expected to ensure that appropriate annual checks had been carried out. For each individual a target HbA1c should be set between 6.5 percent and 7.5 percent based on the risk of macrovascular and microvascular complications (NICE, Management of Blood Glucose, 2002). For the purposes of the Quality Outcomes Framework (QOF) 7.5 (or equivalent) has been selected as an optimal level of control for the purposes of audit and reporting.	items (e.g., retinal screening), but it is the practice's responsibility to ensure that they have been done. This set of indicators relates to both Type 1 and Type 2 diabetes. Although the care of patients with Type 1 diabetes may be shared with specialists, the general practitioner would still be expected to ensure that appropriate annual checks had been carried out. Reaching optimal levels of control (HbA1c 7.5 or less) in diabetic patients is difficult. For this reason a second outcome indicator has been introduced to encourage working with patients with high HbA1c to bring the level to 10 or less. Where fructosamine is used, for example in patients with haemoglobinopathies, local standards may need to be developed for this indicator. The fructosamine value	published guidance on a number of aspects of diabetic control. The indicators for diabetes are generally those which would be expected to be done, or checked in an annual review. There is no requirement on the general practitioner (GP) practice to carry out all these items (e.g., retinal screening), but it is the practice's responsibility to ensure that they have been done. This set of indicators relates to both Type 1 and Type 2 diabetes. Although the care of patients with Type 1 diabetes may be shared with specialists, the general practitioner would still be expected to ensure that appropriate annual checks	published guidance on a number of aspects of diabetic control. The indicators for diabetes are generally those which would be expected to be done, or checked in an annual review. There is no requirement on the general practitioner (GP) practice to carry out all these items (e.g., retinal screening), but it is the practice's responsibility to ensure that they have been done. This set of indicators relates to both Type 1 and Type 2 diabetes. Although the care of patients with Type 1 diabetes may be shared with specialists, the general practitioner would still be expected to ensure that appropriate annual checks

had been carried out. If total cholesterol is greater than 5.0 mmol/l, statin therapy to reduce cholesterol should be initiated and titrated as necessary to reduce total cholesterol to less than 5 mmol/l. There is ongoing debate concerning the intervention levels of serum cholesterol in diabetic patients who do not apparently have cardiovascular disease. Further National Guidance is awaited. The age when a statin should be initiated is unclear. It is pragmatically suggested that all diabetic patients over the age of 40 with a cholesterol of greater than 5 mmol/l should be treated with a	Where fructosamine is used, for example in patients with haemoglobinopathies, local standards may need to be developed for this indicator. The fructosamine value is derived as follows: Fructosamine = (HbA1c – 1.61)/0.017 = 346umol/l The evidence for the targets for HbA1c are based on the Diabetes Control and Complications Trial (DCCT) study in Type 1 diabetes, which found few microvascular complications in those with HbA1c below 7.5 (N Engl J Med, 1993). The authors of the NICE guidelines for Type 2 diabetes (2002) use this to argue for HbA1c levels below 7.5 in Type 2 diabetics. Although there is less direct evidence to support a specific threshold for risk of macrovascular disease in Type 2 diabetes, the 7.5 percent threshold	is derived as follows: Fructosamine = (HbA1c – 1.61)/0.017 = 346umol/l It is recognised that there may be variations in test availability and in normal ranges in different parts of the UK. If this is the case, the primary care organisation (PCO) may stipulate a different but equivalent range for this indicator.	had been carried out. Blood pressure lowering in people with diabetes reduces the risk of macrovascular and microvascular disease. Hypertension in people with diabetes should be treated aggressively with lifestyle modification and drug therapy (SIGN 55, 2001). The most commonly identified target level for blood pressure in patients with diabetes is 140/80. This is the level that health professionals should aim for. A slightly higher level (145/85) is used as the audit standard in common with other indicators.	had been carried out. Weight control in overweight subjects with diabetes is associated with improved glycaemic control. There is little evidence to dictate the frequency of recording but it is general clinical practice that body mass index (BMI) is assessed at least annually.

	statin. Below the age of 40 a decision needs to be reached between the doctor and the patient and may involve assessment of other risk factors and the actual age of the patient.	seems reasonable as a quality indicator for the purposes of QOF, and should play a role in shifting the overall distribution of blood glucose downwards in those with diabetes. It is recognised that there may be variations in test availability and in normal ranges in different parts of the UK. If this is the case, the primary care organisation (PCO) may stipulate a different but equivalent range for this indicator, but it should be noted that the National Diabetes Support Team has advised that all laboratories should now report DCCT aligned results. This issue is discussed in the English NSF under Standards: Supplementary information: Clinical care of adults with diabetes: Monitoring blood glucose control (NSF, 2002).			

Primary Measure Domain	Outcome The validity of measures depends on how they are built. By examining the key building blocks of a measure, you can assess its validity for your purpose. For more information, visit the Measure Validity page.	Outcome The validity of measures depends on how they are built. By examining the key building blocks of a measure, you can assess its validity for your purpose. For more information, visit the Measure Validity page.	Outcome The validity of measures depends on how they are built. By examining the key building blocks of a measure, you can assess its validity for your purpose. For more information, visit the Measure Validity page.	Outcome The validity of measures depends on how they are built. By examining the key building blocks of a measure, you can assess its validity for your purpose. For more information, visit the Measure Validity page.	Process The validity of measures depends on how they are built. By examining the key building blocks of a measure, you can assess its validity for your purpose. For more information, visit the Measure Validity page.
Evidence Supporting the Criterion of Quality	• A formal consensus procedure involving experts in relevant clinical, methodological, and organizational sciences • One or more research studies published in a National Library of Medicine (NLM) indexed, peer-reviewed journal	• A clinical practice guideline or other peer-reviewed synthesis of the clinical evidence • A formal consensus procedure involving experts in relevant clinical, methodological, and organizational sciences • One or more research studies published in a National Library of Medicine (NLM) indexed, peer-reviewed journal	• A formal consensus procedure involving experts in relevant clinical, methodological, and organizational sciences	• A clinical practice guideline or other peer-reviewed synthesis of the clinical evidence • A formal consensus procedure involving experts in relevant clinical, methodological, and organizational sciences	• A clinical practice guideline or other peer-reviewed synthesis of the clinical evidence • A formal consensus procedure involving experts in relevant clinical, methodological, and organizational sciences

Evidence Supporting the Value of Monitoring the Aspect of Population Health	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Evidence Supporting the Value of Monitoring Use of Service	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Extent of Measure Testing	Unspecified	Unspecified	Unspecified	Unspecified	Unspecified
Denominator Inclusions/Exclusions	**Inclusions** Patients with diabetes **Exclusions** Exclude those patients age 16 years and under and patients with gestational diabetes. See "Description of Case Finding" field for exception reporting.	**Inclusions** Patients with diabetes **Exclusions** Exclude those patients age 16 years and under and patients with gestational diabetes. See "Description of Case Finding" field for exception reporting.	**Inclusions** Patients with diabetes **Exclusions** Exclude those patients age 16 years and under and patients with gestational diabetes. See "Description of Case Finding" field for exception reporting.	**Inclusions** Patients with diabetes **Exclusions** Exclude those patients age 16 years and under and patients with gestational diabetes. See "Description of Case Finding" field for exception reporting.	**Inclusions** Patients with diabetes **Exclusions** Exclude those patients age 16 years and under and patients with gestational diabetes. See "Description of Case Finding" field for exception reporting.
Relationship of Denominator to Numerator	All cases in the denominator are equally eligible to appear in the numerator	All cases in the denominator are equally eligible to appear in the numerator	All cases in the denominator are equally eligible to appear in the numerator	All cases in the denominator are equally eligible to appear in the numerator	All cases in the denominator are equally eligible to appear in the numerator
Numerator Inclusions/Exclusions	**Inclusions** Number of patients from the denominator whose last measured total	**Inclusions** Number of patients from the denominator in whom the last HbA1c is 7.5 or less	**Inclusions** Number of patients from the denominator in whom the last HbA1c is 10 or less	**Inclusions** Number of patients from the denominator in whom the last blood pressure is	**Inclusions** Number of patients from the denominator whose notes record body

	cholesterol within the previous 15 months is 5 mmol/l or less **Exclusions** Unspecified	(or equivalent test/reference range depending on local laboratory) in the previous 15 months **Exclusions** Unspecified	(or equivalent test/reference range depending on local laboratory) in the previous 15 months **Exclusions** Unspecified	145/85 or less* ***Note**: The pressure must have been measured in the previous 15 months. **Exclusions** Unspecified	mass index (BMI) in the previous 15 months **Exclusions** Unspecified
Measure Results Under Control of Health Care Professionals, Organizations and/or Policymakers	The measure results are somewhat or substantially under the control of the health care professionals, organizations and/or policymakers to whom the measure applies.	The measure results are somewhat or substantially under the control of the health care professionals, organizations and/or policymakers to whom the measure applies.	The measure results are somewhat or substantially under the control of the health care professionals, organizations and/or policymakers to whom the measure applies.	The measure results are somewhat or substantially under the control of the health care professionals, organizations and/or policymakers to whom the measure applies.	The measure results are somewhat or substantially under the control of the health care professionals, organizations and/or policymakers to whom the measure applies.
Data Source	Laboratory data Medical record Registry data	Laboratory data Medical record Registry data	Laboratory data Medical record Registry data	Medical record Registry data	Medical record Registry data
Level of Determination of Quality	Not Individual Case	Not Individual Case	Not Individual Case	Not Individual Case	Individual Case
Allowance for Patient Factors	Unspecified	Unspecified	Unspecified	Unspecified	Unspecified
Scoring	Rate	Rate	Rate	Rate	Rate
Interpretation of Score	Better quality is associated with a higher score	Better quality is associated with a higher score	Better quality is associated with a higher score	Better quality is associated with a higher score	Better quality is associated with a higher score
Current Use	Internal quality improvement National reporting	Internal quality improvement National reporting	Internal quality improvement National reporting	Internal quality improvement National reporting	Internal quality improvement National reporting

	Pay-for-performance	Pay-for-performance	Pay-for-performance	Pay-for-performance	Pay-for-performance
Care Setting	Physician Group Practices/Clinics	Physician Group Practices/Clinics	Physician Group Practices/Clinics	Physician Group Practices/Clinics	Physician Group Practices/Clinics
Professionals Responsible for Health Care	Physicians	Physicians	Physicians	Physicians	Physicians
Lowest Level of Health Care Delivery Addressed	Group Clinical Practices	Group Clinical Practices	Group Clinical Practices	Group Clinical Practices	Group Clinical Practices

3) List of Cardiologist Measures

Measures

☑ Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for tobacco use. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003776

☑ Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for assessment of exercise capacity. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003783

☑ Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for blood pressure control. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003777

☑ Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for lipid control. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003778

☑ Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for physical activity habits. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003779

Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for assessment of weight management. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003780

☑

Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for diabetes mellitus or impaired fasting glucose. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003781

☑

Acute coronary syndrome (ACS): percent of patients hospitalized with ACS found to be high or moderate-high risk patients with cardiology involvement in care within 24 hours of acute arrival or if AMI as inpatient, within 24 hours of initial ECG or first positive troponin whichever is earlier. Veterans Health Administration 2007 Oct NQMC:002485

☑

Coronary heart disease: the percentage of patients with newly diagnosed angina (diagnosed after 1 April 2003) who are referred for exercise testing and/or specialist assessment. British Medical Association National Health System (NHS) Confederation 2006 Feb NQMC:001878

☑

Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for adherence to preventive medications. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003784

☑

Cardiac rehabilitation: percentage of cardiac rehabilitation program(s) in the healthcare system that meet the specified performance measure criteria for communication with healthcare providers. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003785

☑

☑ Cardiac rehabilitation: percentage of cardiac rehabilitation programs in the health system that meet this specified performance measure criteria for monitoring response to therapy and documenting program effectiveness. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003786

☑ Cardiac rehabilitation: percentage of eligible inpatients with a qualifying event/diagnosis who have been referred to an outpatient cardiac rehabilitation program prior to hospital discharge or have a documented medical or patient-centered reason why such a referral was not made. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003771

☑ Cardiac rehabilitation: percentage of cardiac rehabilitation programs in the healthcare system that meet specified structure-based performance measure criteria. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003773

☑ Cardiac rehabilitation: percentage of cardiac rehabilitation programs in the healthcare system that meet the specified performance measure criteria for assessment of risk for adverse cardiovascular events. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003774

☑ Cardiac rehabilitation: percentage of patients in the healthcare system's cardiac rehabilitation program(s) who meet the specified performance measure criteria for depression. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003782

☑ Cardiac rehabilitation: percentage of patients in an outpatient clinical practice who have had a qualifying event/diagnosis during the previous 12 months, who have been referred to an outpatient cardiac rehabilitation program. American Association of Cardiovascular and Pulmonary Rehabilitation/American College of Cardiology/American Heart Association 2007 Sep NQMC:003772

Appendix B – Example Measures from the MMIC

Criteria	0	1	2	3
Follow up system for abnormal results	None	Policy developed	Staff verbalize understanding	Documentation demonstrates compliance
External diagnostic studies have been read by physician prior to filing in record as evidenced by chart review	<75%	76-85%	86-95%	>95%
Non-compliance with treatment plan addressed	No evidence	Process in place for identification and addressing	Evidence in record that compliance or non-compliance identified	Evidence in record that non-compliance addressed
Medication flow sheet present in chart that includes lab work associated with medication	<50%	50-75%	76-90%	>90%
Methodology in place for tracking trends for chronic disease management	<50%	50-75%	76-90%	>90%
Nationally recognized guidelines are used for patients with diabetes	No		Yes	
System is in place for managing and documenting after hours calls	No	Log system	Notes to primary physician	Primary physician signs off and call is documented in record
Process is in place for evaluating clinical practice of physicians	No			Yes
Patient education system is in place and utilized as demonstrated in medical record	No	<60%	61-90%	>90%
Visit summary given to patient at end of visit including all follow-up instructions	No	<60%	61-90%	>90%

Appendix B: Final Recommendations Report

QUALITY MEASURES FOR PUBLIC REPORTING:
FINAL RECOMMENDATIONS TO THE MINNESOTA DEPARTMENT OF HEALTH
February 6, 2009

BACKGROUND

Through MN Community Measurement (MNCM), the Minnesota health care community has pioneered collaborative health care reporting: building a set of measures that have become both more sophisticated and less administratively burdensome; establishing a process that allows for the collection of quality measure data from medical groups as well as health plans; and providing for the reporting of Minnesota quality data to health care providers and to consumers. Now MN Community Measurement has contracted with the Minnesota Department of Health (MDH) to assist the state in establishing a unified statewide quality reporting system for health care providers. In turn, MN Community Measurement is working with community partners including Stratis Health, the Minnesota Medical Association, the Minnesota Hospital Association and the University of Minnesota School of Public Health to assist us in completing this work.

In December 2008, MN Community Measurement completed an inventory of measures in use across the country for public reporting of quality information. The inventory of measures was presented to MDH and at a series of public meetings. That measurement inventory serves as the basis from which recommendations regarding a subset of measures recommended for initial inclusion in the statewide quality reporting system were made.

To meet the implementation timeline set forth in Minnesota's 2008 health reform law and the statutory requirement that the statewide quality reporting system be initially based on existing quality indicators, MDH and MN Community Measurement agreed that the first set of recommendations for quality measures would draw on measures already in use on a voluntary basis in Minnesota. The existing measures for ambulatory care settings, described further below, resulted from the collaborative efforts of health care providers, health plans, MN Community Measurement, the Minnesota Hospital Association (MHA), and Stratis Health in previous years. The Commissioner of the Minnesota Department of Health will take MN Community Measurement's recommendations under advisement in determining which measures will be chosen for initial inclusion in the statewide quality reporting system. Data for this first round of measures will be collected in 2009 and publicly reported in 2010.

The identification and establishment of measurement priorities as well as the development of the specific measures for inclusion in the statewide quality reporting system will be an iterative process. New science or updated measurement methodologies may necessitate changes to previously included measures or for the replacement of existing measures with others of the same type (e.g., a primary care measure publicly reported in 2010 could be replaced with a different primary care measure in 2011). Therefore, MN Community Measurement's contract with MDH calls for MNCM to repeat the measurement identification and development process in each year of the contract (i.e., 2010, 2011, and 2012). In each of those three years, we will utilize a community-driven process, including extensive formal and informal participation from stakeholders to make preliminary recommendations on measures for public reporting to MDH, solicit public comments, and then refine our recommendations based on stakeholder input.

The community-wide collaborative measure development process established by MN Community Measurement is multi-staged and occurs over several years. Generally, two years are needed for measure development before public reporting can begin; one year is needed to develop a new measure and one year is needed for voluntary data collection and voluntary public reporting to ensure that providers have appropriate systems in place to collect the necessary information. Finally, the measurement is available for public reporting on all data providers during the third year.

The multi-stage development process for the measures that will be included in the statewide quality reporting system is outlined in the chart below. The measures shown in the chart are the minimum number of measures that will be added to the statewide quality reporting system in each year of the contract. The different colors are used to show how a measure works its way through the multi-stage development process over time.

Measurement Development Process Stages

	Stage 1: Determine and select future measurement priorities; Develop new measures	Stage 2: Data collection on new measures begins; Voluntary data submission; Voluntary public reporting	Stage 3: Public Reporting Statewide
First Year: 2009 Calendar Year	1 Primary care measure 1 Specialty care measure 5 New hospital measures supported by clinical-data enhanced database[1]	Depression measure – primary care Depression measure – specialty care Health information technology (HIT) measures Patient experience measures	Existing MNCM and Minnesota Hospital Quality Report[2] measures 12 Additional AHRQ[3] inpatient hospital measures
Second Year: 2010 Calendar Year	1 Additional primary care measure 2 Additional specialty care measures Continuation from first year of development of 5 new hospital measures supported by clinical-data enhanced database[1]	1 Primary care measure 1 Specialty care measure	Previous year's measures, plus: Depression measure – primary care Depression measure – specialty care HIT measures Patient experience measures 2 Additional AHRQ[3] inpatient hospital measures
Third Year: 2011 Calendar Year	2 Additional specialty care measures	1 Additional primary care measure 2 Additional specialty care measures	Previous year's measures, plus: 1 Primary care measure 1 Specialty care measure 5 New hospital measures supported by clinical-data enhanced database[1] 2 Additional AHRQ[3] inpatient hospital measures
Fourth Year: January 1, 2012 – July 1, 2012		2 Additional specialty care measures	Previous year's measures, plus: 1 Additional primary care measure 2 Additional specialty care measures 2 Additional AHRQ[3] inpatient hospital measures

[1] Clinical-data enhanced database will integrate clinical data with administrative data; [2] The Minnesota Hospital Quality Report can be seen at www.mn.hospitalquality.org; [3] AHRQ is the Agency for Healthcare Research and Quality

COMMUNITY-DRIVEN PROCESS

Over the years, MNCM has relied on its Reporting Advisory Committee (RAC) to review and make recommendations on new measures and reporting policies, and to test ideas and strategies related to data collection for ambulatory care settings. The RAC is comprised of physicians and other clinicians, health plan representatives, and technical specialists. For purposes of MNCM's Public Reporting and Payment Incentive contract with MDH to lead a broad-based stakeholder-informed process to develop the statewide quality reporting system and quality incentive payment system, two consumer representatives will also be added to the RAC. When selecting new measures, the RAC utilizes criteria that have been adapted from those used by the National Quality Forum (NQF); these criteria are discussed below. Measure recommendations made by the RAC are informed by national measures where available, as well as through consultation with the Institute for Clinical Systems Improvement (ICSI) on the most recent guidelines approved by Minnesota stakeholders based on a review of national research and evidence. In addition, the RAC also relies on subcommittees of content experts when evaluating particular measurement areas.

The Hospital Quality Reporting Steering Committee will serve a parallel role to the RAC relative to the development and recommendation of inpatient hospital measures to be included in the statewide quality reporting system. The Hospital Quality Reporting Steering Committee is currently under development and will be convened in the coming months for this work. Like the RAC, the Steering Committee will be comprised of broad-based community stakeholders.

In addition to the establishment of the statewide quality reporting system, Minnesota's 2008 health reform legislation also requires the development of a quality incentive payment system. Utilizing a community-driven process, MN Community Measurement will work to develop recommendations for measures that will be included in the quality incentive payment system for both ambulatory care settings and hospitals as well as the proposed methodology for the incentive-based payment system. This process will involve the Incentive Payment Work Group and the Hospital Quality Reporting Steering Committee. The Work Group and Steering Committee will be charged with providing feedback on and advising changes where appropriate on recommendations for the quality incentive payment system made by the MNCM-lead partnership. Like the RAC and Steering Committee, the membership of the Work Group will also include a broad base of relevant stakeholder groups.

MEASURES OF AMBULATORY CARE

CRITERIA FOR RECOMMENDATIONS

The MN Community Measurement Reporting Advisory Committee (RAC) considered the following criteria in making these recommendations regarding which measures should be part of the statewide quality reporting system:

> **Degree of impact** – the magnitude of the individual and societal burden imposed by a clinical condition, including disability, mortality and economic costs.

> **Degree of improvability** – the extent of the gap between current practices and evidenced-based practices (variation) and the likelihood that the gap can be closed and conditions improved through changes in the clinical processes, as well as the opportunity to achieve improvement in the six quality aims laid out by the Institute of Medicine in their March 2001 report titled *Crossing the Quality Chasm: A New Health System for the 21st Century (a brief of the report can be found at* http://www.iom.edu/Object.File/Master/27/184/Chasm-8pager.pdf).

> **Degree of inclusiveness** – the relevance of a measure to a broad range of individuals with regard to (a) age, gender, socioeconomic status, and race/ethnicity; (b) the generalizability of quality improvement strategies across the spectrum of health care conditions; and (c) the capacity for change across a range of health care settings and providers.

> **National consensus** – the measure has either been developed or accepted/approved through a national consensus effort (e.g., National Quality Forum or Physician Consortium for Performance Improvement).

> **Degree of performance variation** – the performance rates show a wide degree of variation (e.g., range from low performer to top performer) from one reported entity to another.

Although the RAC considered the above criteria when it deliberated on the measures currently in use, they may not have explicitly discussed each element of the criteria and the ways in which it was or was not addressed by a particular measure and we did not document the discussion in that way. Therefore, while MN Community Measurement can and will take care to explicitly discuss each element of the criteria when making recommendations for new measures, we cannot provide full documentation of those criteria for existing measures.

RECOMMENDATIONS FOR MEASURES

MN Community Measurement recommends that, in order to get a quick start, the state's quality reporting system include the measures currently in use on a voluntary basis by Minnesota's health care providers and health plans:

- **Optimal Diabetes Care** - the percentage of patients with diabetes (Types 1 and 2) ages 18-75 who reached all five treatment goals:

  – HbA1c <8*

  – Blood Pressure <130/80

  – Low Density Lipoprotein (LDL) <100

  – Daily Aspirin Use

  – Documented Tobacco Free

  *Approved by MNCM RAC and Board in Feb. 2009

Rationale:

Degree of Impact: *As noted by MDH in its Health Fact Sheet on Diabetes in Minnesota (*http://www.health.state.mn.us/diabetes/FactSheet2008.pdf*), diabetes is a high impact clinical condition. One in four Minnesotans either have diabetes or are at high risk of developing it. Diabetes is the sixth leading cause of death in Minnesota and is a leading cause of cardiovascular disease and stroke (risk is 2-4 times higher in people with diabetes), non-traumatic lower extremity amputations (13 times higher risk among Medicare beneficiaries with diabetes than Medicare beneficiaries without diabetes), blindness among people 20 – 74 years of age (diabetes is the leading cause) and end-stage renal disease (leading cause). Diabetes costs Minnesota $2.7 billion annually, including medical care, lost productivity and premature mortality.*

Degree of Improvability: *The best evidence of improvability for optimal diabetes care is the degree of improvement from year to year. In 2007, the average level of optimal diabetes care was 14%, in 2008 it was 17%.*

Degree of Inclusiveness: *There is evidence, according to MDH, that diabetes is relevant to a broad range of individuals with regard to race/ethnicity: of new cases of End-Stage Renal Disease (of which diabetes is the leading cause), 75 cases per million are among non-Hispanic whites, while 433 cases per million are among Native Americans, 314 cases per million are among Hispanics, 177*

cases per million are among Asian-Americans, and 105 cases per million are among African-Americans.

National Consensus: *Our measure matches the guidelines developed by the Institute for Clinical Systems Improvement (ICSI) and we are seeking NQF endorsement for our measure in 2009.*

Degree of Performance Variation: *In 2008, four years after we began reporting on this measure, at some clinics only 3% of the patients reached the five treatment goals that make up this all-or-none composite measure, while at other clinics 33% did so.*

- **Optimal Vascular Care** - the percentage of patients with vascular disease ages 18-75 who reached all four treatment goals:

    - Blood Pressure < 130/80

    - LDL <100

    - Daily Aspirin Use

    - Documented Tobacco Free

Rationale:

Degree of Impact*: According to MDH, vascular disease is a high impact clinical condition (see MDH Fact Sheets on Heart Disease in Minnesota and Stroke in Minnesota at http://www.health.state.mn.us/divs/hpcd/chp/cvh/documents/2007mnheartdiseasefactsheet.pdf and http://www.health.state.mn.us/divs/hpcd/chp/cvh/documents/2007mnstrokefactsheet.pdf). Over 21% of all deaths in Minnesota are due to heart disease and over 7% are due to stroke, making them the second and third leading causes of death, respectively, in the state behind cancer. Heart disease patients incur over $827M and stroke patients over $61M in direct hospital costs alone in Minnesota every year, according to MDH.*

Degree of Improvability*: We cannot look at average improvement over last year in achieving optimal vascular care, because our measure specifications changed from 2007. We can assume that because there is significant variation (see below) that there is an opportunity for improvement.*

Degree of Inclusiveness: *In Minnesota, according to MDH, Native American men experience a 66% higher heart disease death rate than white men, and Native American women have a 33% higher death rate than white women. In Minnesota, the stroke death rate is 34% higher in African-Americans, 28% higher in Native Americans, and 11% higher in Asian-Americans compared to*

whites. Also according to MDH, heart disease is more prevalent in men than in women and men are at greater risk before age 65.

National Consensus: *Our vascular care measure matches ICSI guidelines and is endorsed by the NQF.*

Degree of Performance Variation: *In 2008, at some clinics only 7% of the patients reached the four treatment goals that make up this all-or-none composite measure, while at other clinics 62% did so.*

- **Use of Appropriate Medicines for Asthma** - percentage of patients ages 5-56 with persistent asthma who were appropriately prescribed medication

Rationale:

Degree of Impact*: According to MDH, asthma is a high-impact condition (see MDH Fact Sheet on Asthma in Minnesota at*

http://www.health.state.mn.us/asthma/documents/factasthmaspring07.pdf),

with estimated costs of $208.6M in hospitalizations, emergency department visits, office visits and medications, and $155M in indirect costs of lost school and work days in 2003. One in fifteen Minnesota children has asthma, and one in twelve Minnesota adults report that they currently have asthma.

Degree of Improvability*: Since we began reporting this measure in 2004, medical groups have improved performance on this measure from an average of 74% of patients prescribed appropriately in 2004 to an average of 92% prescribed appropriately in 2008.*

Degree of Inclusiveness: *According to the Asthma and Allergy Association of America "ethnic differences in asthma prevalence, morbidity and mortality are highly correlated with poverty, urban air quality, indoor allergens, and lack of patient education and inadequate medical care." Whether for socioeconomic reasons or urban air quality, adults living in the Minneapolis-St. Paul metro area are more likely to have asthma than residents of Greater Minnesota.*

National Consensus: *Our asthma care measure matches ICSI guidelines and is endorsed by the NQF.*

Degree of Performance Variation: *In 2008, in some medical groups 81% of the patients were prescribed asthma medications appropriately, while in others 100% did so.*

- **Appropriate Treatment for Children with Upper Respiratory Infection** - percentage of children ages three months to 18 years with diagnosis of URI who were <u>not</u> given antibiotic within three days of episode

Rationale:

*<u>**Degree of Impact:**</u> According to NCQA's State of Health Care Quality Report (available on-line at http://www.ncqa.org/tabid/836/Default.aspx), $227M is spent annually on inappropriate treatment for the common cold. The common cold is most often viral, not treatable by antibiotics. Yet, nearly one-quarter of children under the age of 15 who visits the doctor's office for a cold receives a prescription for an antibiotic. Inappropriate use of antibiotics increases drug resistance and increases the individual's risk of becoming infected with a drug-resistance infection.*

*<u>**Degree of Improvability**</u>: This measure has not shown a high degree of improvability as measured by average improvement over time: in 2006, the average rate of appropriate treatment was 86%, in 2007 it was 84% and in 2008 it was again 86%.*

*<u>**Degree of Inclusiveness:**</u> No data.*

*<u>**National Consensus:**</u> Our measure of appropriate care for children with upper respiratory infections matches ICSI guidelines and is endorsed by the NQF.*

*<u>**Degree of Performance Variation:**</u> In 2008, in some medical groups only 16% of children were treated appropriately for upper respiratory infections (meaning they were not given an antibiotic prescription within three days) while in others 97% were.*

- **Appropriate Testing for Children with Pharyngitis** - percentage of children ages 2-18 years with sore throats who were given an antibiotic and a group A strep test for episode period.*Rationale:*

*<u>**Degree of Impact:**</u> According to NCQA's State of Health Care Quality Report (available on-line at http://www.ncqa.org/tabid/836/Default.aspx), pharyngitis, or sore throat, is most commonly caused by viruses. While antibiotics are needed to treat bacterial pharyngitis, they are not useful for treating viral pharyngitis. Only 25 to 50 percent of sore throat cases in children are caused by Group A streptococcus bacteria, more commonly referred to as strep throat. Before antibiotics are prescribed, a simple diagnostic test is necessary to validate bacterial origin of a sore throat. Unfortunately, a diagnostic test is not always performed before antibiotics are prescribed. Inappropriate use of antibiotics increases drug resistance and increases the individual's risk of becoming infected with a drug-resistance infection. NCQA reports that one study found that in 36*

percent of cases where a patient received antibiotics and underwent a test for strep throat, the test came back negative.

Degree of Improvability*: This measure has shown some degree of improvability as measured by average improvement over time: in 2006, the average rate of appropriate testing was 82%, in 2007 it was 81% and in 2008 it was up to 85%.*

Degree of Inclusiveness: *No data.*

National Consensus: *Our measure of appropriate testing for children with pharyngitis matches ICSI guidelines and is endorsed by the NQF.*

Degree of Performance Variation: *In 2008, in some medical groups no children with pharyngitis were tested appropriately while in others 99% were.*

- **Breast Cancer Screening** -percentage of women ages 42-69* who had mammogram in past 2 years

- **Cervical Cancer Screening** - percentage of women ages 24-64 who received one or more Pap tests in past 3 years

- **Colorectal Cancer Screening** - percentage of adults ages 51-80 who had appropriate colorectal cancer screenings

- **Cancer Screening Combined** - percentage of adults ages 51-80 who received appropriate cancer screening services (breast, cervical, colorectal)

  * This will be the age range reported in 2009, changed from the currently reported age range of 32-69.

Rationale:

Degree of Impact: *According to MDH, cancer is the leading cause of death in Minnesota for persons between the ages of 35 and 74. Each year, an estimated 20,600 Minnesotans are diagnosed with cancer and 9,000 die from the disease. (see*
http://www.health.state.mn.us/divs/hpcd/ccs/info/disparit.pdf)

NCQA reports that "mammography screening for women 50 to 69 years of age can reduce breast cancer mortality by up to 35 percent through early detection." NCQA also reports that "Early detection (of cervical cancer) is critical. Cervical cancer rarely causes pain or noticeable symptoms until it is so advanced that it is unresponsive to treatment. Cervical cancer has a five-year survival

rate of more than 90 percent when the cancer is localized, but only 13 percent once the cancer has spread throughout the body." On colorectal cancer screening, NCQA reports that "although symptoms are uncommon in early- stage colorectal cancer, treatment at this stage is extremely effective, with a five-year survival rate of more than 90 percent. Symptoms rarely appear until later stages of the disease, at which point the patient's chances of survival decrease substantially."

Degree of Improvability: *These measures have shown some degree of improvability as measured by average improvement over time. For example: in 2006, the average rate of appropriate colorectal cancer screening was 58%, while in 2008 it was up to 63%. The average rate at which adults ages 51 – 80 received all appropriate cancer screenings went up from 47% in 2006 to 54% in 2008.*

Degree of Inclusiveness: *According to MDH, the breast cancer mortality rate in Minnesota is 50% higher in black women than in white non-Hispanic women even though the incidence rates are similar. MDH notes that "a greater proportion of black women have their cancers diagnosed at a later, less treatable stage." For cervical cancer, black women in Minnesota have an incidence rate that is four times as high as the rate for white women, and Native American and Asian American women have a rate three times as high as the rate for white women.*

National Consensus: *Our measures of cancer screening match ICSI guidelines and are endorsed by the NQF.*

Degree of Performance Variation: *In 2008, in some medical groups 44% of patients received appropriate cervical cancer screenings while in others 100% did; in some medical groups 47% received appropriate breast cancer screenings while in others 100% did; in some medical groups 25% of patients received appropriate colorectal cancer screening while in others 91% did.*

- **Chlamydia Screening** - percentage of sexually active women ages 16-24* who had at least one test for chlamydia infection.

  *This will be the age range reported in the 2009 report, changed from the currently reported age range of 16-25.

Rationale:

Degree of Impact: *According to MDH: "Chlamydia infection is the most commonly reported sexually transmitted disease (STD) in Minnesota. In 2007, 13,412 chlamydia cases (273 per 100,000 population) were reported, representing a 4% increase from 2006. NCQA reports that "left untreated, chlamydia can cause permanent damage to a women's fallopian tubes, uterus and surrounding tissue. Other effects of chlamydia include urethritis, cervicitis, pelvic inflammatory*

disease (PID), infertility, ectopic pregnancy or chronic pelvic pain. Women that are pregnant and have a chlamydial infection are at higher risk for miscarriage, a premature rupture of membranes, preterm labor, low birth weight and infant mortality. 20 to 25 percent of newborns exposed to their mother's chlamydia develop chlamydial conjunctivitis.

Degree of Improvability*: These measures have shown a significant degree of improvability as measured by average improvement over time: in 2004, only 29% of female patients had at least one test for Chlamydia during the year, but by 2008 44% did.*

Degree of Inclusiveness: *According to the MDH Disease Control Newsletter, adolescents and young adults are at highest risk for acquiring chlamydial infection. The chlamydia rate is highest among 20 to 24-year-olds (1,592 per 100,000), with the next highest rate among 15 to 19-year-olds (1,071 per 100,000). The incidence of chlamydia among adults 25 to 29 years of age (716 per 100,000) is considerably lower but has increased in recent years. The chlamydia rate among females (390 per 100,000) is more than twice the rate among males (153 per 100,000). This difference is likely due to more frequent screening among women.*

The incidence of chlamydia infection is highest in communities of color. The rate among blacks (1,871 per 100,000) is over 14 times higher than the rate among whites (130 per 100,000). Although blacks comprise approximately 4% of Minnesota's population, they account for 28% of reported chlamydia cases. Rates among Asian/Pacific Islanders (311 per 100,000), American Indians (504 per 100,000), and Hispanics (646 per 100,000) are two to five times higher than the rate among whites.

Chlamydia infections occur throughout the state, with the highest reported rates in Minneapolis (769 per 100,000) and St. Paul (659 per 100,000). In 2007, the greatest increases for chlamydia were seen in the suburbs and Greater Minnesota with increases of 4% and 8%, respectively."

National Consensus: *Our measure of Chlamydia screening matches ICSI guidelines and is endorsed by the NQF.*

Degree of Performance Variation: *There is a high degree of performance variation between medical groups on this measure: in some medical groups, only 4% of women received appropriate screening, while in others 77% did.*

- **Childhood Immunization** - percentage of children two years of age who had appropriate shots by second birthday

Rationale:

Degree of Impact: *NCQA reports that "Childhood immunizations are responsible for the control of many infectious diseases once common in the U.S., including polio, measles, diphtheria, pertussis (whooping cough), rubella (German measles), mumps, tetanus, and Haemophilus influenzae type b (Hib).Prior to routine vaccination, hepatitis B infected 24,000 infants and children each year. Three out of every 1,000 people who contract measles die of the disease. Measles is one of the most infectious diseases in the world. More than 90 percent of people who are not immune will get the virus if they are exposed to it. Every dollar spent on Hib vaccine saves $3.40. Every dollar spent on hepatitis B vaccine saves $3.60. Every dollar spent on varicella vaccine saves $6.30.6*

NCQA further reports that "childhood immunization through DTaP, Td, Hib, IPV, MMR, hepatitis B and varicella vaccines save $9.9 billion in direct medical costs and $43.3 billion in indirect costs (such as time away from work to care for sick children) in one year. A child with chicken pox misses an average five to six days of school; adult caretakers miss an average three to four days of work.

One-third of lifelong hepatitis B virus infections, which can lead to liver failure and death, result from infections in infants and young children. If the measles vaccine was discontinued in the U.S., three to four million measles cases would occur annually and result in more than 1,800 deaths. Discontinuing Hib immunization would result in approximately 20,000 cases per year of invasive disease and 600 deaths."

Degree of Improvability: *These measures have shown a significant degree of improvability as measured by average improvement over time: in 2006, 52% of children received all their appropriate immunizations, but by 2008 77% did.*

Degree of Inclusiveness: *According to the Centers for Disease Control (http://www.cdc.gov/datastatistics/2007/childimmunization/) "there continue to be small racial/ethnic differences in the percentage of 19- to-35-month-old children receiving the recommended vaccination series. Children who live below the poverty level are less likely to be vaccinated than children who live at or above the poverty level. Because a substantial percentage of black children lived below the poverty level, coverage for black children overall was low compared with white children. Therefore, even though the 2006 survey found that black, non-Hispanic children had lower vaccination rates than white, non-Hispanic children for the series of routine vaccines, the difference was likely related to socioeconomic status and household income rather than race."*

National Consensus: *Our measure of childhood immunizations matches ICSI guidelines and is endorsed by the NQF.*

Degree of Performance Variation: *There is variation between medical groups on this measure: in some medical groups, 45% of children received all of their vaccines by their second birthday, while in others 93% did.*

The MN Community Measurement RAC also recommends the following as new measures:

- **Depression measure, primary care** –
  - Six Month Remission Rate (PHQ-9 score <5 at six months); outcome measure demonstrating improved mental health for patients with depression
  - Use of the PHQ-9 Tool (patient has a PHQ-9 done at least once during the time frame); process measure to track use of new tool used for diagnosis, treatment and monitoring depression care
  - Collected through Direct Data Submission
- **Depression measure, behavioral health specialists** – Includes patients with primary depression diagnosis
  - Six Month Remission Rate (PHQ-9 score <5 at six months); outcome measure demonstrating improved mental health for patients with depression
  - Use of the PHQ-9 Tool (patient has a PHQ-9 done at least once during the time frame); process measure to track use of new tool used for diagnosis, treatment and monitoring depression care
  - Collected through Direct Data Submission

Rationale:

Degree of Impact: *NCQA reports that "direct treatment of depression accounts for only $12.4 billion—about 28 percent—of its total treatment cost. Lost productivity and absenteeism account for the remainder: $44 billion. Depression may lead to appetite and sleep disturbances, anxiety, irritability, decreased concentration, and greatly increases the risk of suicide. The overall health bills of those who suffer from depression are 70 percent higher than those who do not. A patient who discontinues antidepressant treatment within six months incurs an average of more than $400 per year in higher medical costs than adherent patients."*

Degree of Improvability: *According to ICSI, "Major depression is a treatable cause of pain, suffering, disability and death, yet primary care providers detect major depression in only 1/3 to 1/2*

of their patients with major depression." This leaves significant room for performance

improvement.

Degree of Inclusiveness: *NCQA reports that "about 13 million American adults suffer from*

depression each year; 1 in 7 Americans will suffer from a major depressive disorder in their lifetime.

While depression affects people of all ages—the median age of onset is 32— depression is especially

prevalent among the elderly. About 1 in 8 people over the age of 65 suffer from depression."

National Consensus: *Our depression measure matches ICSI guidelines and we are seeking NQF*

endorsement in 2009.

Degree of Performance Variation: *No data. We are currently conducting voluntary collection of*

data from clinics, which will allow us, among other things, to determine the degree of variation

among those clinics.

- **Health information technology –**
    - Self-reported medical group survey assessing their use of HIT
    - As stated in IOM report, the use of IS has potential to improve each of the 6 aims of the
      health care system by helping clinicians manage large amounts of clinical information
    - Report available in mid 2009

Rationale:

Degree of Impact: *According to the NQF, "adoption of HIT by clinicians has been shown to reduce*

medical errors by increasing access to information thereby improving response times to abnormal

results, eliminating repetitive testing and providing clinical decision-support tools to facilitate

evidence-based care. Evidence has shown a decrease in medication errors by up to 20 percent and a

decrease in per admission costs by more than 12 percent when clinicians use HIT."

Degree of Improvability: *The Minnesota e-Health Information Technology Adoption Status*

(http://www.health.state.mn.us/e-health/hitassessmentsummary2008.pdf) compiled data

available about the adoption of HIT in various health care settings, although it did not collect

information about the purposes for which the HIT was being used. Among them, the finding in a

survey conducted by the American Hospital Association (AHA) in2006, which assessed availability of

EHRs. In Minnesota, it found "96% of respondents are actively considering, testing or using IT for

clinical purposes. Among the respondents, 9% have fully implemented EHR, 58% have partially

implemented EHR and 29% have no EHRs yet." Another was a 2006 survey by the Consortium of

Rural Health Research Centers Survey which assessed HIT adoption in Critical Access Hospitals: "the survey found that nationally, and in Minnesota, CAHs have relatively high use rates for administrative and financial HIT applications, but much lower use rates for a number of clinical applications. Only 23% of the responding CAHs are using EHRs, and only 21% were using prescriber order entry."

Degree of Inclusiveness: *No data.*

National Consensus: *Our survey tool incorporates the nine voluntary consensus standards adopted by the NQF in August 2008.*

Degree of Performance Variation: *No data.*

- **Patient experience** -
  - Using national CG-CAHPS survey; four domains:
    - Getting Appointments & Health Care When Needed
    - How Well Doctors Communicate
    - Courteous and Helpful Office Staff
    - Overall Rating
  - Surveys administered by medical groups (vendors) using MNCM specifications
  - First pilot report in early 2009

Rationale:

Degree of Impact: *No data.*

Degree of Improvability: *No data.*

Degree of Inclusiveness: *No data.*

National Consensus: *This measures addresses a sixth Aim of the Institute of Medicine – patient centeredness.*

Degree of Performance Variation: *No data.*

In addition, the RAC recommends:

- **Lead Screening**
  - The percentage of children 2 years of age who had one or more capillary or venous lead blood tests for lead poisoning by their second birthday
  - Relevance to MN Health Care Programs

- Medical group performance variation exists

- HEDIS hybrid method measure collected by health plans

Rationale:

Degree of Impact: *According to NCQA, "an estimated 310,000 children in the U.S. remain at risk for exposure to harmful levels of lead. Very high levels of lead exposure may result in serious, long-term neurological conditions or even death. Elevated blood lead levels are associated with an estimated $1,300 in avoidable medical costs per child and an estimated $3,300 in avoidable special education costs. Based on the reduction in lead exposure since the 1970s, the estimated increase in earnings for children two years of age in 2000 is between $110 billion and $319 billion over their lifetimes."*

Degree of Improvability: *According to NCQA, only six in ten children are screened for lead poisoning.*

Degree of Inclusiveness: *No data.*

National Consensus: *Our lead screening measure matches ICSI guidelines and is a NQF endorsed measure.*

Degree of Performance Variation: *To determine whether there was performance variation, we reviewed initial health plan data, and found that there was significant performance variation, from a low of 45% of children screened to a high of 98%.*

- **Appropriate Management of Adult Acute Bronchitis**
    - The percentage of adults 18-64 years of age with a diagnosis of acute bronchitis who were not dispensed an antibiotic prescription
    - An overuse measure – a higher rate indicates appropriate treatment of adults with bronchitis (i.e., the proportion for whom antibiotics were not dispensed)
    - HEDIS administrative method measure collected by health plans

Rationale:

Degree of Impact: *According to NCQA's State of Health Care Quality Report (available on-line at http://www.ncqa.org/tabid/836/Default.aspx), "between 65 and 80 percent of patients with acute bronchitis receive an antibiotic despite evidence that, with few exceptions, they are ineffective." Inappropriate use of antibiotics increases drug resistance and increases the individual's risk of becoming infected with a drug-resistance infection.*

Degree of Improvability: *No data.*

__Degree of Inclusiveness:__ Elderly patients are particularly likely to receive unnecessary antibiotics.

__National Consensus:__ Our measure of avoidance of antibiotic treatment in adults with acute bronchitis matches ICSI guidelines and is endorsed by the NQF.

__Degree of Performance Variation:__ No data.

DATA SPECIFICATIONS

The data specifications for all recommended measures, including measures currently in use and new measures, are provided in the Appendix of this document.

MEASURES OF HOSPITAL CARE

CURRENTLY REPORTED HOSPITAL MEASURES

There are a number of hospital-specific performance measures that are currently publicly reported that could be considered as candidates for the measures that would serve as the basis for payment incentive systems. The measurement of hospital quality and performance has been largely driven by national requirements, especially by CMS (Centers for Medicare & Medicaid Services) and the Joint Commission, through a national system of data collection; and our work will build on the national efforts to be as aligned as possible, as described here.

For Minnesota's measurement and reporting efforts, it is important to note that there are two distinct types of acute care hospitals in the state, as defined by their Medicare payment and reimbursement system: 53 PPS hospitals (Prospective Payment System, which are generally the medium and large hospitals) and 79 CAH (Critical Access Hospitals, which are small, rural hospitals). Minnesota has one of the largest numbers of CAHs in the state, both in terms of the number of hospitals, and in the number of hospitals per capita. A Minnesota quality measurement system ideally should assess the quality of care at both types of hospitals.

PPS hospitals currently collect and publicly report a set of 40 measures, and do so under a national Medicare "Pay for Reporting" program that allows these hospitals to earn their full Annual Payment Update in exchange for reporting these measures:

- Heart Attack (AMI) Care: 8 measures (7 process measures, plus mortality rates)

- Heart Failure Care: 5 measures (4 process measures, plus mortality rates)

- Pneumonia Care: 8 measures (7 process measures, plus mortality rates)

- Surgical Care: 7 measures (all process measures)

- Children's Asthma: 2 measures (both process measures)

- Experience of Care: 10 measures derived from HCAHPS patient survey

Critical Access Hospitals are not subject to the same "Pay for Reporting" program, and only some of the above listed measures are relevant for the scope of services provided by CAHs. Specifically, the Heart Failure (6 measures) and Pneumonia (8 measures) are relevant and appropriate for small rural hospitals. In Minnesota, 68 of the 79 CAHs currently collect and report on at least one of these measures,

demonstrating their commitment to quality and transparency, even though they do not have financial incentives for reporting.

There is a national data repository for data reporting that both PPS and CAHs utilize for the above measures. The process measures are all collected quarterly from medical records (i.e., through chart abstraction) by the hospitals, and are subject to validation checks by CMS, supported by the Medicare QIO (in Minnesota, Stratis Health). The results are posted and updated quarterly on the Hospital Compare web site, at www.hospitalcompare.hhs.gov.

In 2005, Stratis Health and the Minnesota Hospital Association launched a Minnesota-specific companion web site to the national Hospital Compare web site, drawing on the data that hospitals already collect and submit, and using these data to calculate an all-or-none Appropriate Care Measure. The Appropriate Care Measure is an easy-to-understand, consumer-friendly measure that indicates whether, for each patient, they received all of the care that they should have given their condition. The Minnesota Hospital Quality Report can be accessed at http://www.mnhospitalquality.org.

In addition, in Minnesota, all hospitals, both PPS and CAH, are required to report to the state when an Adverse Event occurs. The Minnesota law was passed in 2003, requiring that hospitals (and outpatient surgical centers and community behavioral health hospitals) report whenever one of 28 such events occur. This information is released publicly, by hospital name, in January each year.

Finally, In addition to the public reporting programs described above, hospitals have a variety of other voluntary data collection and reporting programs, and data reporting that is done about them, including: Leapfrog patient safety measures (derived from survey response), AHRQ (Agency for Healthcare Research and Quality) Quality Indicators (derived from claims data), and more.

Here is the list of measures reported on the Minnesota Hospital Quality Report (MHQR):

- **Heart attack**:

    - Patients given aspirin at arrival

    - (†)Patients given eta locker at arrival

    - Patients given ACE inhibitor or ARB for left ventricular systolic dysfunction (LVSD)

    - Patients given thrombolytic medication within 30 minutes of arrival

    - Patients given PCI within 120 minutes of arrival

    - Patients given smoking cessation advice/counseling

    - Patients given aspirin at discharge

- Patients given beta blocker at discharge

- (*)Appropriate care measure

- **Heart failure**:

  - Patients given assessment of left ventricular function (LVF)

  - Patients given ACE inhibitor or ARB for left ventricular systolic dysfunction (LVSD)

  - Patients given smoking cessation advice/counseling

  - Patients given discharge instructions (LVF)

  - (*)Appropriate care measure

- **Pneumonia**:

  - (†)Patients given oxygenation assessment

  - Patients having a blood culture performed prior to first antibiotic received in hospital

  - Patients given initial antibiotic(s) within 4 hours after arrival

  - Patients given the most appropriate initial antibiotic(s)

  - Patients assessed and given pneumococcal vaccination

  - Patients given smoking cessation advice/counseling

  - (*)Appropriate care measure

- **Surgical care**:

  - Surgery patients who received preventative antibiotic(s) one hour before incision

  - Surgery patients whose preventative antibiotic(s) are stopped within 24 hours after surgery

  - Prophylactic antibiotic selection for surgical patients

  - Surgery patients with recommended venous thromboembolism prophylaxis ordered

  - Surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours prior to surgery to 24 hours after surgery

- **Survey of patient's hospital experiences:**

- Communication with doctors

- Communication with nurses

- Responsiveness of hospital staff

- Pain control

- Communication about medicines

- Cleanliness of hospital environment

- Quietness of hospital environment

- Discharge instructions

- Overall rating of the hospital

- Willingness to recommend the hospital to others

- **Hospital acquired infection**:

  - Cardiac surgery patients with controlled 6 A.M. postoperative blood glucose

  - Surgery patients with appropriate hair removal

  - Ventilator bundle compliance

  - Central line bundle compliance

  - Surgical site infection rate for vaginal hysterectomy

  - Surgical site infection rate for total knee arthroplasty

  (†)Measure will be discontinued

  (*)Appropriate care measure is a composite, "all or none" measure combining all of the individual measures within the care category

Most of the data for the MHQR is based on all-payer chart-abstracted data submitted to CMS. Most Minnesota hospitals participate in the CMS project, including all of the acute care hospitals over 25 beds. Even hospitals under 25 beds mostly participate, but for them the CMS program has no financial consequences like it does for the larger ones. The new infection measures are mandatory for all Minnesota hospitals, however.

RECOMMENDATIONS FOR NEW MEASURES

The team working on the hospital reporting aspects of this project -- Stratis Health and Minnesota Hospital Association, under the leadership of MN Community Measurement -- recommends a 3-part measurement and reporting strategy for 2009:

1. Include as part of Minnesota's measurement and reporting program the 25 current clinical process measures already being collected and reported by Minnesota hospitals in AMI, heart failure, pneumonia, surgical care, starting in 2009.

2. As required by the contract with the State, include as part of Minnesota's measurement and reporting program the new 12 AHRQ patient safety indicators starting in 2009.

3. Implement a process through which additional new hospital measures for public reporting can be identified, developed and recommended, by July 2009 (through a parallel process to the utilized for the development of new ambulatory measures). The recommendations will include measures relevant to rural hospital care and scope of services, for implementation in 2010, with the option of continuing the process to identify innovative measures, especially those that begin to reflect care coordination across ambulatory and acute care.

These recommendations would produce 38 measures of hospital quality available to the state and the public by the end of 2009, of which 15 are relevant to small rural hospitals; and would support a process through which additional new measures would be identified for data collection in 2010 and public reporting in 2011.

As mentioned earlier, in an effort to get a quick start on implementing the statewide quality reporting system for hospital care, MDH and MN Community Measurement agreed that twelve specific measures from the Agency for Healthcare Research and Quality (AHRQ) would be recommended for initial inclusion in the statewide quality reporting system. MN Community Measurement subcontracted with the Minnesota Hospital Association to fulfill this portion of its Public Reporting and Payment Incentive contract with MDH.

The AHRQ Quality Indicators are measures of health care quality that make use of readily available hospital inpatient administrative data. They are organized in four modules:

- Inpatient Quality Indicators (28 provider level measures)

- Patient Safety Indicators (20)

- Prevention Quality Indicators (0)

- Pediatric Quality Indicators (13 – newest – pediatric version of patient safety indicators, mostly)

Other states, including Colorado and Texas, use AHRQ measures for public reporting.

CRITERIA FOR RECOMMENDATIONS

In order to select 12 measures out of the roughly 50 available AHRQ measures in its Inpatient Quality Indicators (IQI) and Patient Safety Indicators (PSI), several factors were considered:

➢ Alignment with other public reporting or quality improvement activities. For example, does the measure relate to prevention of adverse health events or to process measures reported to the Centers for Medicare and Medicaid Services?

➢ Number of hospitals with significant volume. Does this apply to most hospitals?

➢ Likelihood of consumer interest. Does this relate to relatively common conditions or procedures?

➢ Coding/severity adjustment issues. Is performance on this measure affected significantly by the accuracy and completeness of coding? *The severity of illness calculation is dependent on accurate and complete coding of secondary diagnoses – this can be inconsistent between facilities and even between coders at the same facility.* Is there some controversy whether the severity adjustment methodology is adequately robust for this measure? Some measures may not do a good job of adjusting for the age of a patient, for example.

➢ Outcome measures. Does the indicator capture the contract's stated preference for measuring performance on outcomes?

AHRQ MEASURES

To meet the condensed timelines specified by the 2008 health reform legislation, a group of experts reviewed the AHRQ measures against these criteria. In the future, this work will be done by the Hospital Quality Reporting Steering Committee. Based on the expert feedback received, we recommend the following measures for public reporting:

● Abdominal aortic aneurysm repair (AAA) – IQI 4
● AAA repair mortality rate – IQI 11
● Coronary artery bypass graft (CABG) – IQI 5
● CABG mortality rate – IQI 12

- Percutaneous transluminal coronary angioplasty (PTCA) – IQI 6
- PTCA mortality rate – IQI 30

Rationale:

These measures align with Leapfrog Initiative measures. Being that these are the most common types of cardiac surgery, they are likely to be of interest to consumers. The volume measures have no coding issues – the measures are defined by the presence of the corresponding procedure code and because these are significant types of surgeries, they will virtually always be coded. Higher volume is also considered a marker for higher quality. Mortality rates are outcome measures, although the severity adjustment is imperfect. Even if the coding for all hospitals were complete and accurate, the secondary diagnoses only explain part of the variance in performance – richer clinical data, like lab values, are needed to more accurately ascertain the patient's severity of illness.

- Hip fracture mortality rate – IQI 19

Rationale:

This measure is a CMS measure. It applies to all hospitals and is meaningful to consumers although of low occurrence. Iit is an outcome measure. The severity adjustment is imperfect, however.

- Decubitus Ulcer – PSI 3

Rationale:

This measure aligns with both a CMS measure and is a state Adverse Health Event measure. It applies to all hospitals. It is of consumer interest as an avoidable condition and is an outcome measure. However, it is subject to coding variations, particularly with regard to whether it is present on admission.

- Death among surgical patients with treatable serious complications – PSI 4

Rationale:

This measure aligns with a CMS measure and is related to reported Adverse Health Events. It applies to and is tracked by most hospitals. It is of consumer interest as an avoidable event and is an

outcome measure. It does have some coding issues as coding of the pertinent complications is often
not uniform across hospitals.

- Post-operative pulmonary embolism or deep vein thrombosis – PSI 12

Rationale:

This measure aligns with the Hospital Quality Alliance Venous Thromboembolism topic and applies
to most hospitals. It may not be of extremely high interest to consumers but does apply to all
surgeries. It is an outcome measure and has fewer coding issues.

- Obstetric trauma (3rd and 4th degree lacerations) – vaginal delivery with instrument – PSI 18
- Obstetric trauma (3rd and 4th degree lacerations) – vaginal delivery without instrument – PSI 19

Rationale:

These measures align with a JCAHO measure reported by some hospitals. It applies to most
hospitals. It will be of interest to consumers as one of the few obstetrical measures available. It is an
outcome measure. However, there are some coding issues and some uncertainty about how these
events can be prevented.

While meeting all of the criteria for every measure would have been ideal, it should be noted that most
do not meet every criteria listed. However, each measure chosen had significant positive attributes that
outweighed the drawbacks relative to other candidate measures. Below follows a discussion of why
other ARQH measures were not chosen:

- **Mortality for specific medical conditions (6 out of 7 indicators not chosen)**

Rationale:
The severity adjustment for these measures is less robust than for surgical conditions – while
theoretically, statewide observed mortality rates should be close to equal to the expected mortality
rates after severity adjustment is applied, in practice, the observed mortality rates for medical
patients who are very old (85+) tend to be higher than expected. In practical terms, this means that
hospitals that treat an older population of patients may show higher than expected mortality rates
for conditions like pneumonia. These indicators may be useful for internal quality improvement

activities, however using them to compare relative performance could lead to erroneous conclusions. Mortality for hip fracture was chosen instead.

- **Mortality for specific surgical conditions (5 out 8 indicators not chosen)**

Rationale:

The three measures chosen – AAA Repair, CABG, and PTCA – measure mortality for surgical conditions that occur at a higher volume and have corresponding volume indicators.

- **Utilization measures (none of 7 chosen)**

Rationale:

The utilization measures are confusing to readers because favorable performance is undefined. For instance, there is controversy in the clinical community about what is the best practice with regard to c-sections and VBACs so these measures could not be easily tied to quality improvement efforts.

- **Volume measures (3 of 6 not chosen)**

Rationale:

The higher volume procedures were chosen. Carotid Endarterectomy was a candidate, along with its mortality measure , but it is lower volume and not all hospitals perform this procedure.

- **Other Patient Safety Indicators**

Rationale:

Some of these have a very low occurrence (<1 per 1000), while others have coding issues. For example, for accidental puncture/laceration, there may be a great deal of individual coder judgment in determining whether a puncture or laceration led to additional treatment, which is the central question of whether it gets coded. Therefore, there may be both under and over reporting of the diagnoses that cause patients to be listed in the numerator. An over-reporter, for example, might list minor punctures that probably had no bearing on the overall treatment. Hospitals that over-report may erroneously appear to have performance issues, and the opposite is true for under-reporting.

- **Composite measures**

  Rationale:

  Though the idea of composite measures has some appeal in that it allows hospitals with lower volume to aggregate their individual measures, the methodology to do this compounds all of the coding and severity issues that are present at the individual measure level. Again, hospitals that tend to treat older populations tend to do worse in these types of mortality measures. CMS does propose to publicly report the following composite measures in the future: death in medical conditions, death in surgical conditions, and overall patient safety.

- **Pediatric measures**

  Rationale:

   These measures are of very low occurrence.

DATA SPECIFICATIONS

The data specifications for the recommended all AHRQ measures are provided in the Appendix.

RURAL-SENSITIVE MEASURES

Given the prevalence of small rural hospitals in Minnesota, there is a need to publicly report hospital measures that are relevant to most hospitals in Minnesota, and to include measures that are not exclusive to only those hospitals with high-volume or that provide inpatient services and surgeries not typically performed in smaller rural hospitals. While some of the currently collected and proposed AHRQ measures are relevant for rural hospitals, several measures relevant to smaller rural hospitals have been developed and are currently being used by hospitals in Minnesota. Measures such as those that focus on care in the emergency room or the experience of care are relevant to all hospitals regardless or volume or services provided and are aligned with other public reporting and quality improvement activities, and will be considered in the process for identifying and recommending new measures by July 2009. Potential measures for consideration include:

- Emergency Department measures: Selecting from a set of 5 timeliness measures (from the recently announced CMS Outpatient Prospective Payment System measures, which are NQF endorsed) for patients presenting with chest pain that is likely to be a heart attack/AMI, and/or from a set of 3 NQF-endorsed measures of time to transfer/admissions for patients with a

variety of conditions.

- Medication Safety: Selecting from a set of 34 measures in 7 domains of safe medication practices (as studied by the Center for Rural Health Policy Analysis).

- Experience of Care: Developing from the HCAHPS patient experience of care survey a composite measure(s) of patient assessment of quality of care.

LOW-VOLUME HOSPITALS

It is our intent to publish all of the hospital data, even hospitals with low volumes. Currently, the data for low volume hospitals are displayed in both the Minnesota Hospital Quality Report and Medicare's HospitalCompare. However, both sites do not list the data for low volume hospitals alongside the high volume hospitals – additional clicks are necessary to access the low volume hospital data. This is a subject that should be addressed within the discussion of how data will be publicly displayed. Following is a table showing how many hospitals each of the selected AHRQ measures affect, as well as how many Critical Access Hospitals are affected:

AHRQ Measures	All Hospitals w/ cases	All Hospitals w/ >10 cases	CAHs w/ cases	CAHs w/ >10 cases
AAA Repair (vol & mort)	26	13	0	0
CABG (vol & mort)	14	14	0	0
PTCA (vol & mort)	17	16	0	0
Hip Fracture mort	90	58	41	13
Decubitus Ulcer	129	124	76	72
Deaths, surgery pts w/ treatable compl	89	45	38	5
Postop PE/DVT	121	108	67	55
Birth Trauma w/ Instrument	91	68	44	22
Birth Trauma w/o Instrument	105	96	57	49

Appendix C: Final Recommendations Report

Preliminary Recommendations of Measures and Methodologies and for Minnesota's Quality Incentive Program

Submitted by the University of Minnesota to Minnesota Community Measurement

February 13, 2009

<u>Introduction</u>

In previous work towards the development of a state quality measurement and payment incentive system, MN Community Measurement, through its subcontract with the University of Minnesota, has inventoried quality measures in use for public reporting and/or for payment incentive programs. In this document, the University of Minnesota, under subcontract with MN Community Measurement, has identified a subset of quality measures as the basis for a quality incentive payment system. This report presents these preliminary recommendations on measures and also some preliminary recommendations on quality incentive methods.

The preliminary recommendations address:

- Which subset of measures from the quality measures identified under Task One that we recommend as the basis for a quality incentive payment system.

- A rationale for why we recommend those measures, as well as a rationale for why the other measures identified in Task One are not recommended as the basis for a quality incentive payment system.

- In selecting a subset of measures, we will maximize the feasible use of outcome-related measures that improve care and lower costs for high volumes of people; focus on chronic conditions; and minimize both providers' administrative burden and duplication of related activities.

- What type of quality incentive methods appear to be most favorable for the state.

<u>Background</u>

Through MN Community Measurement, the Minnesota health care community has pioneered collaborative health care reporting: building a set of measures that have become both more sophisticated and less administratively burdensome; establishing a process that allow for the collection of quality measure data from medical groups as well as health plans; and providing for the reporting of Minnesota quality data to health care providers and to consumers. Now MN Community Measurement has contracted with the Minnesota Department of Health (MDH) to assist the state in establishing a unified statewide quality reporting system for health care providers. In turn, MN Community Measurement is working with community partners including Stratis Health, the Minnesota Medical Association, the Minnesota Hospital Association and the University of Minnesota School of Public Health to assist us in completing this work.

In December 2008, MN Community Measurement completed an inventory of measures in use across the country for public reporting of quality information. On February 6, 2009, Minnesota Community Measurement, along with community partners including Stratis Health, the Minnesota Medical Association, the Minnesota Hospital Association and the University of Minnesota, recommended a set of quality measures for public reporting for the State of Minnesota.

111

There are two important planning considerations in play:

- First, MDH does not envision selecting measures for public reporting or for incentive payments as a one-time process; new science or updated measurement methodologies may call for changes in previously existing measures or for replacement of existing measures with others of the same type.
- Second, MDH recognizes that the process that this health care community has followed to develop and implement measures in a collaborative way, takes time. Generally, it takes a year to develop a new measure, a year of voluntary data collection and voluntary public reporting to make sure, among other things, that health providers have appropriate systems in place to collect the information, and then the measurement information is ready for public reporting and perhaps use in awarding incentive payments on all data providers in the third year.

Over time, the Reporting Advisory Committee, comprising physicians and other clinicians, purchasers, consumers, technical specialists and health plans, worked together to establish priorities for these recommended measures. When considering new measures, the Reporting Advisory Committee used criteria that have been adapted from the National Quality Forum.

The criteria included:
1. Degree of impact
2. Degree of improvability
3. Degree of inclusiveness
4. National consensus
5. Degree of performance variation

The measures that were recommended on February 6, 2009 by the Minnesota Community Measurement are the following:

Ambulatory Care Measures

1. Optimal Diabetes Care - the percentage of patients with diabetes (Types 1 and 2) ages 18-75 who reached all five treatment goals:

    a. HbA1c <7

    b. Blood Pressure <130/80

    c. Low Density Lipoprotein (LDL) <100

    d. Daily Aspirin Use

    e. Documented Tobacco Free

2. Optimal Vascular Care - the percentage of patients with vascular disease ages 18-75 who reached all four treatment goals:

    a. Blood Pressure < 130/80

    b. LDL <100

    c. Daily Aspirin Use

    d. Documented Tobacco Free

3. Use of Appropriate Medicines for Asthma - percentage of patients ages 5-56 with persistent asthma who were appropriately prescribed medication

4. Appropriate Treatment for Children with Upper Respiratory Infection - percentage of children ages three months to 18 years with diagnosis of URI who were not given antibiotic within three days of episode

5. Appropriate Testing for Children with Pharyngitis - percentage of children ages 2-18 years with sore throats who were given an antibiotic and a group A strep test for episode period.

6. Breast Cancer Screening -percentage of women ages 42-69* who had mammogram in past 2 years
   * This will be the age range reported in 2009, changed from the currently reported age range of 32-69.

7. Cervical Cancer Screening - percentage of women ages 24-64 who received one or more Pap tests in past 3 years

8. Colorectal Cancer Screening - percentage of adults ages 51-80 who had appropriate colorectal cancer screenings

9. Cancer Screening Combined - percentage of adults ages 51-80 who received appropriate cancer screening services (breast, cervical, colorectal)

10. Chlamydia Screening - percentage of sexually active women ages 16-24* who had at least one test for chlamydia infection.

    *This will be the age range reported in the 2009 report, changed from the currently reported age range of 16-25.

11. Childhood Immunization - percentage of children two years of age who had appropriate shots by second birthday

12. Depression measure, primary care –
    - Six Month Remission Rate (PHQ-9 score <5 at six months); outcome measure demonstrating improved mental health for patients with depression
    - Use of the PHQ-9 Tool (patient has a PHQ-9 done at least once during the time frame); process measure to track use of new tool used for diagnosis, treatment and monitoring depression care
    - Collected through Direct Data Submission

13. Depression measure, behavioral health specialists – Includes patients with primary depression diagnosis
    - Six Month Remission Rate (PHQ-9 score <5 at six months); outcome measure demonstrating improved mental health for patients with depression
    - Use of the PHQ-9 Tool (patient has a PHQ-9 done at least once during the time frame); process measure to track use of new tool used for diagnosis, treatment and monitoring depression care
    - Collected through Direct Data Submission

14. Health information technology –
    - Self-reported medical group survey assessing their use of HIT
    - As stated in IOM report, the use of IS has potential to improve each of the 6 aims of the health care system by helping clinicians manage large amounts of clinical information
    - Report available in mid 2009

15. Patient experience -
    - Using national CG-CAHPS survey; four domains:
      o Getting Appointments & Health Care When Needed
      o How Well Doctors Communicate
      o Courteous and Helpful Office Staff
      o Overall Rating
    - Surveys administered by medical groups (vendors) using MNCM specifications
    - First pilot report in early 2009

16. Lead Screening
    - The percentage of children 2 years of age who had one or more capillary or venous lead blood tests for lead poisoning by their second birthday
    - Relevance to MN Health Care Programs
    - Medical group performance variation exists
    - HEDIS hybrid method measure collected by health plans

17. Appropriate Management of Adult Acute Bronchitis
    - The percentage of adults 18-64 years of age with a diagnosis of acute bronchitis who were <u>not</u> dispensed an antibiotic prescription
    - An overuse measure – a higher rate indicates appropriate treatment of adults with bronchitis (i.e., the proportion for whom antibiotics were not dispensed)
    - HEDIS administrative method measure collected by health plans

Hospital Care Measures

- Abdominal aortic aneurysm repair (AAA) – IQI 4
- AAA repair mortality rate – IQI 11
- Coronary artery bypass graft (CABG) – IQI 5
- CABG mortality rate – IQI 12
- Percutaneous transluminal coronary angioplasty (PTCA) – IQI 6
- PTCA mortality rate – IQI 30
- Hip fracture mortality rate – IQI 19
- Decubitus Ulcer – PSI 3
- Death among surgical patients with treatable serious complications – PSI 4
- Post-operative pulmonary embolism or deep vein thrombosis – PSI 12
- Obstetric trauma (3rd and 4th degree lacerations) – vaginal delivery with instrument – PSI 18
- Obstetric trauma (3rd and 4th degree lacerations) – vaginal delivery without instrument – PSI 19

Table I presents the timeline for the development and implementation of measures for public reporting. Measures used for quality incentives are to be selected from these measures.

Table I: Timeline for Measure Development and Implementation from Task One

	First Year (Calendar Yr 2009)	Second Year (Calendar Yr 2010)	Third Year (Calendar Yr 2011)	Fourth Year (To July 1 2012)
Public Reporting on All Data Providers	• Existing MNCM and Minnesota Hospital Quality Report data as reported at http://www.mnhospitalquality.org/ • 12 additional AHRQ inpatient measures	Previous year plus: • Depression – primary care • Depression – specialists • HIT measure • Patient experience • 2 additional AHRQ inpatient measures	Previous year plus: • 1 Primary care measure • 1 specialty care measure • Five new hospital measures integrating clinical data with administrative data • 2 additional AHRQ	Previous year plus: • 1 additional primary care measure • 2 additional specialty care measures • 2 additional AHRQ inpatient measures

			inpatient measures	
Data Collection on New Measures; Voluntary Data Submission; Voluntary Public Reporting	• Depression measure – primary care • Depression measure – specialists • Health information technology measures • Patient experience measures	• 1 Primary care measure • 1 specialty care measure	• 1 additional primary care measure • 2 additional specialty care measures	• 2 additional specialty care measures
Develop new measure	• 1 Primary care measure • 1 specialty care measure • Clinical-data enhanced database to support five new hospital measures	• 1 additional primary care measure • 2 additional specialty care measures • Clinical-data enhanced database to support five new hospital measures	• 2 additional specialty care measures	

Source: MNCM Quality Measures for Public Reporting: Final Recommendations to the Minnesota Department of Health, February 6, 2009

Measures Currently in Use for Quality Incentive Programs for Minnesota Providers

There are a number of provider quality incentive programs that have been implemented in Minnesota. Appendix B provides tables listing measures that are currently in use for quality incentive programs in Minnesota.

Ambulatory Care Quality Incentive Programs

In general, physician incentive programs are sponsored by individual health plans using plan-specific or MNCM measures. CMS sponsors the Physician Quality Reporting Initiative (PQRI) that began in 2007 and is expanding. The Minnesota Medical Association (MMA) maintains an annually updated inventory of measures that are being used for quality incentive programs that involve Minnesota physicians.

Hospital Quality Incentive Programs

The primary hospital quality incentive program that involves Minnesota hospitals is CMS Hospital Compare. Like the PQRI program, the incentive payment is in the form of a higher annual fee update for reporting measures. This program is expected to expand the number of reportable measures and evolve to a pay-for-performance program over time. The

115

CMS Hospital Compare program currently includes 40 measures. Hospital reported process measures are used as the basis for the financial incentive; but outcomes measures in the form of risk-adjusted mortality rates are also reported by CMS based on administrative data.

These measures include:
- Heart Attack (AMI) Care: 8 measures (7 process measures, plus mortality rates)
- Heart Failure Care: 5 measures (4 process measures, plus mortality rates)
- Pneumonia Care: 8 measures (7 process measures, plus mortality rates)
- Surgical Care: 7 measures (all process measures)
- Children's Asthma: 2 measures (both process measures)
- Experience of Care: 10 measures derived from HCAHPS patient survey

The MNCM February 6 report recommended the universe of quality measures from which to select measures for the initial implementation of quality incentives. Regarding hospital measures, the report indicated that, "There are two distinct types of acute care hospitals in the state, as defined by their Medicare payment and reimbursement system: 53 PPS hospitals (Prospective Payment System, which are generally the medium and large hospitals) and 79 CAH (Critical Access Hospitals, which are small, rural hospitals). Minnesota has one of the largest numbers of CAHs in the state, both in terms of the number of hospitals, and in the number of hospitals per capita. A Minnesota quality measurement system ideally should assess the quality of care at both types of hospitals."

"Critical Access Hospitals are not subject to the same "Pay for Reporting" program, and only some of the above listed measures are relevant for the scope of services provided by CAHs. Specifically, the Heart Failure (6 measures) and Pneumonia (8 measures) are relevant and appropriate for small rural hospitals. In Minnesota, 68 of the 79 CAHs currently collect and report on at least one of these measures, demonstrating their commitment to quality and transparency, even though they do not have financial incentives for reporting."

The following is a list of the specific Hospital Compare Measures (excludes the mortality measures)

Heart attack:
- Patients given aspirin at arrival
- (†)Patients given beta locker at arrival
- Patients given ACE inhibitor or ARB for left ventricular systolic dysfunction (LVSD)
- Patients given thrombolytic medication within 30 minutes of arrival
- Patients given PCI within 120 minutes of arrival
- Patients given smoking cessation advice/counseling
- Patients given aspirin at discharge
- Patients given beta blocker at discharge
- (*)Appropriate care measure

Heart failure:
- Patients given assessment of left ventricular function (LVF)
- Patients given ACE inhibitor or ARB for left ventricular systolic dysfunction (LVSD)
- Patients given smoking cessation advice/counseling
- Patients given discharge instructions (LVF)
- (*)Appropriate care measure

Pneumonia:
- (†)Patients given oxygenation assessment

116

- Patients having a blood culture performed prior to first antibiotic received in hospital
- Patients given initial antibiotic(s) within 4 hours after arrival
- Patients given the most appropriate initial antibiotic(s)
- Patients assessed and given pneumococcal vaccination
- Patients given smoking cessation advice/counseling
- (*)Appropriate care measure

Surgical care:
- Surgery patients who received preventative antibiotic(s) one hour before incision
- Surgery patients whose preventative antibiotic(s) are stopped within 24 hours after surgery
- Prophylactic antibiotic selection for surgical patients
- Surgery patients with recommended venous thromboembolism prophylaxis ordered
- Surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours prior to surgery to 24 hours after surgery

Survey of patient's hospital experiences:
- Communication with doctors
- Communication with nurses
- Responsiveness of hospital staff
- Pain control
- Communication about medicines
- Cleanliness of hospital environment
- Quietness of hospital environment
- Discharge instructions
- Overall rating of the hospital
- Willingness to recommend the hospital to others

Hospital acquired infection:
- Cardiac surgery patients with controlled 6 A.M. postoperative blood glucose
- Surgery patients with appropriate hair removal
- Ventilator bundle compliance
- Central line bundle compliance
- Surgical site infection rate for vaginal hysterectomy
- Surgical site infection rate for total knee arthroplasty

To provide further context on the differences between PPS hospitals and rural hospitals, we quote from the February 6 report, "Measures such as those that focus on care in the emergency room or the experience of care are relevant to all hospitals regardless or volume or services provided and are aligned with other public reporting and quality improvement activities, and will be considered in the process for identifying and recommending new measures by July 2009.

Potential measures for consideration include:

- Emergency Department measures: Selecting from a set of 5 timeliness measures (from the recently announced CMS Outpatient Prospective Payment System measures, which are NQF endorsed) for patients presenting with chest pain that is likely to be a heart attack/AMI, and/or from a set of 3 NQF-endorsed measures of time to transfer/admissions for patients with a variety of conditions.

- Medication Safety: Selecting from a set of 34 measures in 7 domains of safe medication practices (as studied by the Center for Rural Health Policy Analysis).

- Experience of Care: Developing from the HCAHPS patient experience of care survey a composite measure(s) of patient assessment of quality of care."

<u>Preliminary Recommendations for Minnesota Pay for Performance</u>

This section presents preliminary recommendations on measures and methods for the initial pay-for-performance implementation for Minnesota's health care reform. See Appendix A for a description and analysis of quality incentive programs, methods in current use, and experience to-date with these initiatives, including a review of empirical research. These preliminary recommendations relate to a number of design issues that are addressed in depth in Appendix A.

General recommendations

We have observed that the implementation of any new pay-for-performance system is an iterative process that takes place over a number of years. When implementation is to be community or statewide, the need to prioritize the tasks in the initial phase of implementation becomes critical to its early success. For a state like Minnesota, with a well established existing quality performance measurement system at the physician practice level, and with the introduction of the Medicare Hospital Compare program and related public reporting of that performance, we believe that the initial phase should focus on the pay-for-performance methodology and not add to the implementation burden by introducing new performance measures.

In later phases, adjustments to the method will be inevitable and appropriate and we would anticipate that measures will be added.

It seems likely that three or more years will be required to introduce truly advanced comprehensive performance measures and to fully realize the comprehensive objectives of the state for an all-payer, value-based purchasing system.

Our recommendations for the initial phase are predicated on using measures and risk adjusters that are currently in place, with efforts to expand scale and scope taking place in a parallel to this initiation implementation period. The recommendations and related measures for quality assessment and public reporting are presented in the context of differentiating the initial implementation phase from later phases.

In the initial phase we recommend building on current MNCM measures used for public reporting, as well as selected Hospital Compare measures used for pay-for-reporting and public reporting. More specifically, we recommend that the MNCM Diabetes measure, the cardiovascular measure (s), asthma management measure, cancer screening measures, and HIT measures be used for pay-for-performance. We would delay using the depression measures, the overuse of antibiotic measures, and the important patient experience measures now under development.

For hospitals, it seems reasonable to build on the Hospital Compare methodology for pay for reporting and public reporting for those measures that have not "ceilinged" (that is, a substantial portion of hospitals have approached the maximum possible performance level.)

Design Issues and Related Recommendations

 While empirical findings about the implementation and impact of pay-for-performance in general, and for comparing alternative approaches, are currently incomplete and somewhat inconsistent, we can outline issues and make some educated judgments to support designing and implementing a pay-for-performance system.

Measures

Recommendation 1) Use well established performance measures for introducing a statewide program of pay-for-performance.

Rationale: Large scale and community-wide examples of pay-for-performance usually followed or were associated with measures that had already been used for public reporting or the private profiling of providers, or piloted in "dry runs". We think a statewide initiative, such as this, requires that the initial measures be well- established in the community, thus measures of physician practice already implemented by MNCM and supported by health plans, as well as measures already implemented by MN hospitals for public reporting, should be regarded as the logical initial candidates for pay-for-performance, with other measures added after they have been piloted.

Another value of using existing measures is that benchmarks are available to use in translating performance into payment and also to assist purchasers in budgeting for the cost of pay-for-performance.

In addition, one of the most common concerns of providers regarding pay-for-performance initiatives that rely on new measures is the cost of collecting and reporting the data. This will be less of a concern in Minnesota, because of its history of MNCM- reported physician performance and because Minnesota hospitals have their performance in the CMS Hospital Compare program reported by the MHA on the MHA website.

This implies that the State may wish to delay implementing the MNCM depression measures until the Diamond project has progressed. Another example of potentially important measures for pay-for-performance concerns measures of patient experience. In this case, the state may decide that the MNCM measures could be implemented in the near future, after the measurement and reporting process has had more time to evolve.

Type of Measures

Recommendation 2) While initially using a subset of existing MMCM and Hospital Compare measures, the State should develop a plan to increase the use of outcome measures, including patient experience, and also to include more clinical outcomes measures. This will become increasingly feasible as more providers submit their data directly to MNCM for measure construction and reporting.

Rationale: The principle of using well established measures limits the number of possible different measures available for the initial phase of implementation, but experimentation with new measures, and with redesign of existing measures, should be ongoing as a stimulus to performance improvement.

Recommendation 3) In the initial implementation phase, we recommend that the State not include "overuse" performance measures, such as an antibiotic use for URIs, in its pay-for-performance program.

Rationale: There is little experience with overuse measures and it may be controversial to establish a state-sponsored pay for performance program that may appear to pay for withholding a service. This more difficult issue should be addressed in a future round.

Number of measures

Recommendation 4) Begin with a relatively small subset of MNCM and Hospital Compare measures that are already use.

Rationale: If recommendation 1 is adopted, the field of candidate measures is relatively limited for the first phase of implementation. However, there is little evidence to date that focusing rewards on a small set of measures leads to poorer quality in areas not eligible for rewards. Expanding the number of measures over time should be a relatively

short-term programmatic objective, but scaling up a comprehensive value-based purchasing structure will require addressing most clinical care systems and processes in some way.

<u>Methods</u>

Amount of payment

Recommendation 5) The amount of the incentive in the initial phase must be considered in the context of an assessment of the potential impact on provider revenues once the program is scaled up. While determining the actual amount will require financial modeling, in general the research suggests that an initial pilot effort involving even a modest (e.g. 1-3%) proportion of provider revenue may be sufficient, if there is certainty that the program will expand to include multiple payers. In Minnesota, the plan is for the program to start with rewards for serving state employees and then quickly expand to all health plan enrollees.

Rationale: Studies show that even relatively modest rewards initially may be effective if providers know with certainty that the scope of the pay for performance effort, in terms of number of patients and payers involved, will increase in a relatively brief time.

Structure of payment

Recommendation 6) Implement a combination pay for improvement and pay for attainment approach. Improvement targets could be based on a percent improvement comparing the gap between the provider's current performance and either 100% or a community benchmark. Rewards also could be paid to providers for achieving a pre-determined benchmark. This two part system should be fine-tuned over time to ensure that providers continue to have an incentive to improve even after reaching the minimum target level. For example, the top performers and the top improvers could be awarded an additional bonus on a tournament basis.

This approach must be modeled so that the payers' budget management requirements and the desire for the incentive to have an impact on provider behavior are considered in the context of current performance levels and variation in performance in Minnesota.

Aggregation of measures

Recommendation 7) We suggest that the State adopt a point system to weight physician practice scores across all performance measures. Hospital Compare composite methods should be used for hospital scoring.

Rationale: An aggregated scoring system will allow fine tuning of the weighting process in addition to setting targets or benchmarks for specific measures. Because the typical weighting methods all have room for improvement, a hospital and also a physician performance scoring panel of experts should be assembled to review and improve upon current practices used to establish scoring weights. The panel should ensure that the weightings in some way reflect two objectives: 1) maximizing the health of the population and 2) maximizing the efficiency of care.

Accounting for case mix and risk

Recommendation 8) in the initial phase, we suggest that the State conduct a brief review of the current exclusion and inclusion criteria in the MNCM and Hospital Compare measures and determine if any changes would be desirable for their use in structuring measures for the purpose of pay-for-performance. This is the first step in case mix/risk adjustment. It is likely that the existing exclusions and inclusion will be adequate for the initial phase.

In subsequent phases of expansion, the State should consider separate performance scoring for Medicaid, Medicare, and commercial populations if sample size problems can be overcome through expansion in the number of purchasers and payers. If stratification is made based on such proxies, or even on direct measures of social factors, the state may need to adjust the weighting in the scoring system to provide greater incentives for improvement for populations with barriers. The issue for risk adjustment is how to be fair to those being assessed, without inadvertently establishing a policy that accepts lower quality care for populations with barriers.

Rationale: Most current P4P programs to date have not used risk adjustment as part of the construction of measures but have relied on both exclusion from the denominator and/or separate scoring for Medicaid, Medicare, and commercial populations. This probably is because payment typically has been made based on the presence or absence of a particular activity recommended in treatment guidelines. For the most part, one would expect that the treatment would be carried out irrespective of patient characteristics. However, as outcome measures are added to the program, this assumption may become less defensible.

Recommendation 9) The State should study the need for and feasibility of developing advanced statistical risk adjustment methods for three types of risk factors: 1) Co-morbidity /severity 2) Social complexity and 3) Patient behavior and non-adherence with recommendations.

Rationale: We do not currently have reliable data nor well tested risk models for introducing severity as a risk adjuster in many care settings. In addition, we do not have data on some critical social factors that can serve as barriers to providing care or improving outcomes. This will become an increasingly important methodological issue as pay-for-performance increases in scale and scope and as performance improves for the less difficult to treat. This objective should be a key element of the second phase of pay-for-performance implementation.

Recommendation 10) The State should study the application of the exception policies that are in use in the physician pay-performance in the UK, in which MDs can exclude specific patients from their denominator based on some nationally established and auditable reasons (e.g. Called patient more than X times to come for the test, but they refused or did not follow-up; or adhering to this guideline is contraindicated). Whether such an option for risk adjustment would work in Minnesota is an open question; however, the study itself could support the development of a new risk adjustment system based on new data and statistical modeling.

Recommendation 11) The State should develop risk adjustment methods for the new outcomes measures that may be included in the pay-for-performance program in the future. Hospital Compare measures include risk adjustment, and this approach should be adopted for MN as well.

<u>Preliminarily Recommended Measures for Initial Implementation of Quality Incentives</u>

Ambulatory Care

- Optimal Diabetes Care - the percentage of patients with diabetes (Types 1 and 2) ages 18-75 who reached all five treatment goals:

- Optimal Vascular Care - the percentage of patients with vascular disease ages 18-75 who reached all four treatment goals:

- Use of Appropriate Medicines for Asthma - percentage of patients ages 5-56 with persistent asthma who were appropriately prescribed medication

- Appropriate Testing for Children with Pharyngitis - percentage of children ages 2-18 years with sore throats who were given an antibiotic and a group A strep test for episode period.

- Breast Cancer Screening -percentage of women ages 42-69* who had mammogram in past 2 years

121

- Cervical Cancer Screening - percentage of women ages 24-64 who received one or more Pap tests in past 3 years

- Colorectal Cancer Screening - percentage of adults ages 51-80 who had appropriate colorectal cancer screenings

- Cancer Screening Combined - percentage of adults ages 51-80 who received appropriate cancer screening services (breast, cervical, colorectal) * This will be the age range reported in 2009, changed from the currently reported age range of 32-69.

- Chlamydia Screening - percentage of sexually active women ages 16-24* who had at least one test for chlamydia infection.

  *This will be the age range reported in the 2009 report, changed from the currently reported age range of 16-25.

- Childhood Immunization - percentage of children two years of age who had appropriate shots by second birthday

- Health information technology – self-reported medical group survey assessing their use of HIT

- Lead Screening - percentage of children 2 years of age who had one or more capillary or venous lead blood tests for lead poisoning by their second birthday

Recommendations of Measures for Initial Implementation of Quality Incentives for PPS Hospitals

The initial implementation of quality incentives should build on Hospital Compare, where measures meet the following criteria:

- Sufficient variability in performance across Minnesota hospitals, and
- Sufficient sample sizes to apply to a large proportion of PPS hospitals.

Upon review of Minnesota Hospital Compare results, it appears that there will be very few measures that meet these criteria. The state must decide whether it wishes to establish a quality incentive program initially to get the methodology in place and begin the process of expanding measures as they are piloted for public reporting. The innovation initially may be restricted to moving from CMS's pay-for-reporting to actual pay-for-performance.

It should be noted that not-paying for medical errors is also a form of quality incentive that is being actively explored and implemented by purchasers and payers, but was believed to be out of context for this Task.

Recommendations for Rural Hospitals

There is no precedent we know of for pay-for-performance for rural hospitals. It is possible that such examples will be found in a wider search of private health plan programs and possibly for state Medicaid programs.

Because of this lack of history, because the reporting of these measures by rural hospitals has only very recently begun, and since the CAHs are paid on a very different basis than PPS hospitals, we recommend that an action plan be developed with the goal of implementing pay for performance for rural hospitals in either 2010 or 2011, depending on findings related to performance benchmarks and appropriate incentive methods.

Next Steps

These preliminary recommendations, made by the University of Minnesota School of Public Health under subcontract to MN Community Measurement, will be fully vetted by the public and by stakeholders. Written comments can be submitted to comments@mnhealthcare.org until Feb. 28, 2009. Two public meetings will be held in February to solicit input on the proposals. Please see http://www.health.state.mn.us/healthreform/measurement/publicmeetings.html for

information on those public meetings. In addition, MN Community Measurement has established two stakeholder workgroups, one addressing recommendations for hospitals, the other for ambulatory providers. For information on meetings of those workgroups, which are open to public observation, please see http://www.health.state.mn.us/healthreform/measurement/workgroup.html. MNCM's final recommendations to MDH are due March 25, 2009.

Appendix D: Final Recommendations Report

Inventory and Literature Review: Pay–for-Performance Methods and Structure

Prepared for Minnesota Department of Health by
The University of Minnesota, under contract with MN Community Measurement

There is a developing literature on pay-for-performance initiatives directed at health care providers in the United States and other countries (Christianson, Leatherman and Sutherland, 2008a,b; Christianson, Leatherman and Sutherland, 2007). Under these initiatives, providers are paid more if they achieve quality benchmarks or demonstrate improvements in the quality of care they provide. Although not common, some initiatives involve financial penalties for providers that fail to achieve quality targets or demonstrate improvement.

In pay-for-performance initiatives, quality is measured in a variety of ways, including: presence of certain characteristics in the practice environment (e.g. maintenance of a patient registry for people with specific chronic illnesses, or use of electronic medical records); carrying out desirable practice activities (e.g. initiating regular contact with patients who have chronic illnesses); conforming to evidence-based practice recommendations (e.g. for diabetics, meeting recommendations for frequency of blood sugar testing), or achieving goals for biologic measures (e.g., for diabetics, blood sugar levels below a specified target.) Sponsors of P4P initiatives include employer groups, health plans, provider organizations, Medicaid programs, and CMS.

In general, the literature finds that P4P initiatives are associated with better quality (Christianson, Leatherman and Sutherland, 2008a). But, it is difficult to determine if the initiatives caused the observed quality improvements, in part because financial incentives usually are employed as one component of an overall quality improvement strategy. Therefore, any quality gains that are observed could be due to payments, or they could be due to other features of the quality improvement strategy, or both.

There is no consistency in the design and implementation of the P4P initiatives that have been evaluated in the published literature. However, irrespective of the problem of determining the causal relationship between P4P initiatives and quality improvement in "real world" situations, the improvements reported in these evaluations are typically modest. Where quality gains are reported, they most often pertain to a subset of the measures on which payment was based (Christianson, Leatherman and and Sutherland, 2008b). Various explanations have been offered for the absence of major, significant effects associated with P4P initiatives, including:

- The incentive payments have not been large enough to change provider behavior.

- The existence of multiple different P4P initiatives sends confusing signals to providers.

- The wrong measures are being use to assess quality improvement.

- Providers do not have sufficient resources to invest in the major practice changes needed to improve quality.

- Providers face other, more pressing, issues (e.g. installing electronic medical records) that draw their attention away from responding to P4P initiatives (which could be considered a variant of "the payment is not large enough" argument.)

- The current practice environment for providers (especially physicians) leaves little time in a patient visit to take the necessary steps required to achieve a P4P reward; under fee-for-service payment, seeing more patients may be a more "efficient" way to increase practice income (if that is the goal).

- In some P4P schemes, the structure of payment creates uncertainty, which reduces the likelihood that providers will invest in practice changes necessary to receive a P4P award. For instance, if the reward is given to the top X% of providers, some providers may achieve a high level of performance on a P4P metric but may not be rewarded for their efforts.

- Physicians believe P4P programs encourage "check box" medicine and distract them from treating the "whole patient."

Most of these explanations relate to the design and implementation of P4P, not flaws in the basic concept. (Although, some analysts argue that incentive schemes such as P4P have not been particularly successful in non-health care settings in part because they are based on a flawed premise: that individual behavior is motivated primarily by financial incentives as opposed to a variety of other possible influences. For a general discussion, see Gagne and Deci, 2005) Because market conditions and the preferences of providers vary across locations and over time, there is no single, optimal P4P program structure. Instead, the challenge for P4P implementers is to design and carry out a P4P initiative that best fits their situation. This involves carefully considering different design options.

In this document, we identify the decisions that we believe, based on our review of the literature, are likely to be the most important in constructing a P4P initiative. However, the literature very seldom provides clear guidance concerning which decision to make, under which conditions. Among other things, this reflects the relatively small number of published evaluations of "real world" P4P programs from which to draw lessons, and the fact that (as noted above) P4P programs typically are embedded in broader quality improvement strategies, which both affects their impact and makes that impact difficult to determine. Consequently, many aspects of P4P have been the subject of debate, without the development of a clear consensus (e.g see Fisher, 2006; Nelson, 2007).

Structure of Payment

A major issue in the design of any P4P program concerns whether to reward achievement of a predetermined benchmark of performance (e.g. 90% of patients in the practice with illness X receive test in time period T), or to reward a predetermined level of improvement (e.g. achievement of Z percent increase in the proportion of patients in the practice with illness X receiving test, measured

from period T to period T+1.) There are other variations to consider. For instance, rather than rewarding improvement in the percent of patients receiving a test, if the measure is continuous the reward could be based on average amount of improvement (say, in a biologic marker.) A very different approach is to reward the top X percent of performers, using a predetermined formula to divide reward dollars (e.g. the top 5% of performers receive 50 % of the reward dollars, the next 5% receive 25%, and so forth.) This is called a "tournament" approach to compensation, because providers compete against each other for an award. The tournament approach can be applied to improvement as well. A final issue in the structure of payment relates to the use of penalties for low performers who do not improve over period T.

Rewarding the achievement of benchmarks: A major argument in favor of awarding P4P dollars for the achievement of pre-specified benchmarks is that the award process is easy to understand. The target is clear to providers, which should make their planning processes easier. For instance, they know that if they institute a new patient tracking process, and it is effective, they will receive a reward. In contrast, under a tournament approach, the provider may make the investment, perform at a higher level of quality as a result, but receive no reward if other providers perform better. From the provider's perspective, the uncertainty of obtaining the reward could discourage investment in quality improvement processes. However, there are three significant drawbacks associated with using pre-determined benchmarks to distribute award dollars. First, providers who already meet the benchmarks are essentially rewarded for their historical performance. They have little incentive to improve. This may not be a satisfactory result from the standpoint of payers who want their P4P programs to improve quality. Second, depending on where the benchmarks are set, providers who are at the low tail of the performance curve may have little incentive to invest in quality improvement. This would be the case if these providers did not believe they could improve enough to meet the benchmark in any given period. Conceptually, the use of benchmarks is likely to be the most effective if they are set at a level above current performance and if most providers felt that the benchmarks could be achieved through reasonable efforts. Third, benchmarks can be problematic for payers because they complicate budgeting for the P4P initiative. If more providers achieve the benchmarks than predicted, costs could be greater than expected or budgeted. This risk increases in situations where there is not a good historical record of provider performance relative to the benchmark. Then, if the benchmark is set too low, a substantial portion of providers may have met it prior to implementation of the initiative. This problem can be solved by using a tournament approach, where the amount of reward dollars available is determined prior to implementing the P4P initiative. Then there is no risk that the payer will exceed the budgeted P4P amount.

Rewarding improvement: The alternative to using benchmarks for structuring rewards is to allocate award dollars based on percentage or nominal improvement. This has the advantage of providing incentives for low performing providers to invest in quality improvement. In fact, these providers may have a better chance of obtaining rewards than higher quality providers. This may be regarded as desirable if payers place greater weight on raising the quality of care offered by low-performing providers. However, it may not seem fair to other providers who perform better, but receive no award dollars, because their improvement was not as great. This issue is particularly relevant when,

for the measures chosen, some providers already are performing near the maximum and have little or no possibility of earning an award.

Practice Example

There are two striking examples in the literature that highlight the challenges involved in structuring payments using a benchmark approach. In the United States, Rosenthal, et. al. (2005) evaluated a physician P4P program implemented by Pacificare, an HMO serving the western part of the United States (and subsequently acquired by UnitedHealthcare.) Pacificare rewarded performance on 5 ambulatory care quality measures and 5 patient-centered measures of service quality. Benchmarks were established based on the 75th percentile of the 2002 performance of the physician groups. These benchmarks were known to the participating medical groups, which also knew their own performance relative to the benchmarks prior to the initiative. An average medical group with 10,000 Pacificare patients had the potential to earn $270,000 per year, which was equal to 5% of the capitation payment to the group from Pacificare, but under 1% of an average group's total revenues. The program awarded about $3,4 million in bonuses from July, 2003-April, 2004 (27% of the potential bonus payments), and the evaluators found a significant, but modest, quality improvement in one of the three clinical quality measures they examined. Interestingly, according to the evaluators, "Physician groups whose performance payments were above the benchmark at baseline captured 75% of bonus payments" (p. 1792) for the measures they examined. In effect, three quarters of the award money was used to reward past performance of medical groups. However, in a somewhat unexpected finding, groups that had the poorest quality scores at the beginning of the P4P initiative demonstrated the greatest improvement.

The second example concerns the physician P4P initiative in the U.K. (Roland, 2004). This program committed up to $3.2 billion in new funds over three years to reward general practitioners for performance relative to 146 quality indicators. Physicians earned points for percentages of patients in their practices meeting predetermined benchmarks. A 75% achievement of benchmarks overall was predicted when setting the budget for the initiative, but in the first year physicians achieved 96.7% of the points for quality indicators (Doran, et. al., 2006). Most of the budgeted monies for this three year program were dispersed in the program's first year. As a result, Campbell, et. al. (2007) observed that "The size of the gains in quality in relation to the costs of pay for performance remains a political issue in the United Kingdom, and the government now accepts that it paid more than it had expected to pay for the improvements in performance" (p. 189). In fact, the government did not have reliable data to use in setting the benchmarks and apparently underestimated existing quality levels. It is likely that a significant portion of the funds actually rewarded practices for their historical performance. Because there were not good baseline measures of quality, it is unclear whether or not the P4P program actually resulted in significant quality improvements in primary care in the U.K.

Amount of Payment

A second major issue relates to the amount of payment necessary to achieve improvement. This can be thought of in terms of a specific dollar amount or a percent of practice income. The "right amount"

necessary to achieve the desired result likely varies with the structure of the payment and how the payment, if received at the physician group level, is used (e.g. included in incentive payments to physicians or physician practices or used of improving group infrastructure, etc.)

Practice Example

There is almost no research that addresses the level of payment needed to achieve desired results in a P4P program. To do so would require a study design in which several P4P programs were compared that were identical except for differences in reward levels. The closest that any published study comes to this design is the comparative analysis of five Medicaid programs conducted by Felt-Lisk, et. al., (2007). However, because the Medicaid programs varied along dimensions other than size of payment, this evaluation could conclude only that the greatest response occurred in the program that offered the largest rewards.

The literature does highlight the considerable variation in potential (and, in some cases, actual) rewards to be found in P4P programs. The U.K. P4P program offered the most generous rewards found in any of the P4P programs where evaluations have been published. At its inception, that program offered the potential for physicians to increase their practice incomes by $77,000 per physician (Roland, 2004); in practice during its first year the program increased the income of general practitioners by an average of $40,000, a considerable percentage increase over their average income of $122,000 to $131,000 before the program was put in place (Doran, et. al., 2006). As noted above, the potential increase in income under Pacificare's P4P program was less than 1% of a medical group's annual revenues.

Most published evaluations do not provide detailed information on the size of the reward in the P4P program, either in absolute or in relative terms. Where this information is provided, the reward typically is relatively small. For instance, the maximum reward for top-performing hospitals in the CMS/Premier P4P demonstration was a 2% increase in Medicare reimbursements, which is an increase of 1% or less in total revenues for most hospitals (Lindenauer, et. al., 2007). In early experiments conducted by Hillman, et. al. (1998, 1999), the bonus for top performing physician practices was 10 percent of a practice's typical capitation payment from the payer. In commenting on the incentive's lack of impact, the authors note that this payer was one of many for these practices. Kouides et. al. (1998) evaluated the impact of paying physicians $0.80 per influenza immunization if their practice immunization rates exceeded 70% and $1.60 if they exceeded 85%. Immunization rates improved by a greater amount in the incentive group, relative to a group of physicians that did not receive incentives, even though Kouides, et. al. (1998) characterized the incentives as "modest." In a P4P program implemented by a health plan in Hawaii, physicians received an average bonus payment of 3.5% for attaining predetermined benchmarks, and physicians who showed significant improvement in scores received a bonus payment of $3,000. In a health plan-sponsored program aimed at diabetes care improvement, the annual distribution of payments ranged from $6,000 to $18,000 (Curtin, et. al., 2006). Larson, Cannon, and Towner (2003) report a relatively small incentive of from 0.5 to 1 per cent of physician income for improvement of care along several dimensions, with half of the payment going towards rewarding improvements in diabetes care.

As these examples illustrate, the level of the awards found in P4P programs varies enormously. In some cases, it appears that reward amounts were set in negotiations with participating physicians, and in other cases they seem to have been set unilaterally by payers. Irrespective of how they were set, "real world" P4P programs are not structured in a way that allows researchers to answer a question that is of fundamental interest to payers: "how much is enough" to generate significant improvements in quality.

Type of Measure

A third design feature concerns the types of measures to use when rewarding performance. The basic issue is whether to use "process of care" measures or "outcome" measures, or some combination of the two. The choice has implications for the cost of data collection and measure construction (and, therefore, the cost-effectiveness of P4P). There are varying arguments. For instance, treatment processes are under the control of providers, whereas outcome measures may be influenced by patient behaviors that providers may not be able to control. Thus, from a provider perspective, process measures may be regarded as a "fairer" basis for rewarding provider performance. But, from a payer perspective, the goal of a P4P initiative ultimately is to improve patient health, so payers may favor P4P programs based on outcome measures.

Practice Example

Most measures used in P4P programs sponsored by health plans in the United States are constructed using claims data (e.g. see Chung, et. al., 2003), and therefore they are process of care measures. Generally, they try to capture the conformance of care with widely-accepted evidence-based treatment guidelines (e.g. see Felt-Lisk, et. al., 2007; Greene, et. al., 2004). They tend to focus on preventive actions relating to screening (e.g. see Armour, et. al., 2004; Langham, Gillam, and Thorogood, 1995; Rosenthal, et. al., 2005) and receipt of immunizations (e.g. Morrow, Gooding and Clark, 1995), and on treatment of chronic illnesses where there are widely accepted medical standards. More indicators are available for diabetes that for any other chronic illness (e.g. see Beaulieu and Horrigan, 2005; Curtin, et. al., 2006; Young, et. al., 2007). Treatment of heart conditions is another area of care that receives considerable attention in P4P initiatives (e.g. see Glickman, et. al., 2007; Nahra, et. al., 2006). Some health plans have included measures that go beyond items for which claims are processed (e.g. delivery of smoking cessation advice). These measures typically require documentation in the medical records of patients, and performance is assessed based on a random sample of patient records. Some P4P programs use a variety of different measures, with the P4P program in the U.K. providing the best example of this. In this program, rewards are given for certain practice characteristics and for patient ratings, along with standard process of care indicators. In general, payers appear to be expanding the number of outcome measures used in their P4P programs, especially intermediate outcome measures such as lipid levels, blood pressure readings and HbA1c levels. These data are available in patient medical records and typically are collected and submitted by the physician practice. The increased use of electronic medical records facilitates this approach, but constructing these measures remains more expensive for providers, and for payers as well if the data are audited. Measures based on patient reports are the most expensive to construct and also are the least common types of measures presently

used in P4P programs in the United States. Five of the 10 measures used by Pacificare in its P4P program (see above) were patient-reported measures of service quality (Rosenthal, et. al., 2005)

Number of Measures

Determining the number of measures to use, irrespective of the type of measure, is an important design decision in any P4P program. The argument for using a large number of measures is that it encourages overall improvement in quality. The argument against it is that the incentive to improve in any one area is weak. Advocates of fewer measures say this can focus provider resources on areas where improvement is needed the most. However, a contrary view is that providers may focus too strongly on the targets, and quality may decrease in areas not included in the P4P program. This concern is no different from that sometimes expressed about paying teachers or schools based on student performance on specific standardized tests. The issue is whether, in the presence of rewards, teachers will "teach to the test;" that is, focus on subjects that will be covered in the test, to the detriment of student learning in other areas.

 Practice example

Active P4P programs vary widely in the number of P4P measures they use. For example, the P4P program aimed at general practitioners in the U.K uses 146 quality indicators while programs sponsored by health plans in the U.S. typically use 10 or fewer measures designed to reflect well-established best practices in preventive care and in the treatment of some chronic illnesses. For instance, researchers have evaluated the impact of several different P4P initiatives addressing only diabetes care. Research relating to the number of measures used in P4P programs has focused primarily on documenting the impact of P4P on un-rewarded aspects of quality.

Glickman, et al. (2007) assessed the impact of the CMS/Premier hospital P4P initiative (described above) on AMI process of care measures. Six measures were included in the P4P program, and the authors tracked these measures, as well as 8 other AMI treatment quality indicators not rewarded under the program. They found significant improvements in two P4P measures where no change in physician practice was required, and the cost of change for hospitals was relatively low (aspirin at discharge and smoking cessation counseling), but no improvement in a composite performance measure that included all 6 P4P metrics. The authors also found no effect—negative or positive—on any of the measures not included in the CMS/P4P initiative. Beaulieu and Horrigan (2005) evaluated a physician P4P program for diabetes care sponsored by a health plan in the United States. Based on data from a small number of physician practices, they reported improvement in 5 of 6 diabetes measures in the P4P program and no impact on quality of care in areas not targeted by P4P for rewards. In evaluating the impact of the U.K.'s P4P program on 42 primary care practices, Campbell, et. al. (2007) focused on measures of quality for heart disease, asthma and type 2 diabetes. They found significant, but modest, improvements in asthma and diabetes quality of care indicators, and no negative impacts on other measures of quality not incented by the P4P program.

In summary, concerns that P4P programs could have a negative impact on quality in areas of care where financial incentives were not applied are not supported by existing research. There are (at

least) two explanations for this research finding. First, there is little credible evidence in evaluations of P4P programs to date that these programs have resulted in substantial improvements in the indicators that they have targeted with financial incentives. It would seem most reasonable to expect negative impacts on other areas of quality in situations where P4P programs had an impact on their targeted measures. Second, it may be that significant resource shifts within practices are not necessary to secure P4P rewards related to many quality measures, so that non-targeted areas of care are not put "at risk" by P4P programs.

Aggregation of Measures

P4P programs that use multiple measures of performance must decide how they will use information from these measures to pay providers. The basic choice is whether to pay providers based on their scores on an aggregate measure of performance or to pay separately for performance on individual measures, with the total payment equal to the sum of these individual payments. When paying providers based on an aggregate score, P4P implementers must decide how to combine the individual scores; that is, they must decide the weight to give to each component in creating the aggregate measure. We could find no discussion of the strengths and limitations of these different approaches in the literature. It seems reasonable to propose that, if an aggregate measure is constructed, the aggregation process should be transparent to providers participating in the P4P program. Providers seeking to achieve rewards for improving quality or achieving threshold levels of quality should have a clear understanding of which aspects of quality are valued the most by the payer. It also is important for payers to understand that a simple adding up of achievement on different measures to form the aggregate score implicitly means that all improvements are valued equally.

Practice Example

The Pacificare P4P program used 10 different measures of quality; 5 are clinical measures and 5 are measures of patient satisfaction. Performance on each measure was rewarded separately, so that the medical group's total reward was the sum of these individual rewards. In contrast, in the CMS/Premier hospital demonstration payment to hospitals was made based on an aggregate measure. According to Lindenauer, et. al (2007), "For each of the clinical conditions, hospitals performing in the top decile on a composite measure of quality for a given year received a 2% bonus payment in addition to the usual Medicare reimbursement rate" (p. 488). The Premier website describes the aggregation process in detail, with examples. Hospitals begin by submitting their raw data to Premier, which calculates a quality index for each clinical area included in the program. The overall index consists of a "Composite Process Rate" and a "Risk-Adjusted Outcomes Index." To calculate the Composite Process Rate, the numerator values for each of the process measures are summed to create a composite numerator, with a composite denominator calculated in the same way. The composite numerator is divided by the composite denominator to generate the overall Composite Process Rate. For outcomes measures, a hospital's actual outcomes rate is divided by its risk-adjusted rate and the result is multiplied by 100. A final Composite Quality Index in each area is calculated by weighting each score within each area equally. The scores are ranked, and hospitals in the top decile or second decile receive

a predetermined payment for their performance in that area. This process is repeated for every clinical area included in the program.

In the U.K P4P initiative, a different number of points was awarded for different performance indicators. Martin Roland, MD, one of the architects of the U.K.'s program, states that "Family practitioners can now earn up to 1000 'points' for achievement in relation to the complex set of indicators that make up the Quality and Outcomes Framework" (Roland, 2004). As an example, for patients with heart disease, if blood pressure has been recorded in the previous 15 months for 25% of patients, the practice receives 1 point. If it has been recorded for 90% or more of patients, the practice receives 7 points. Points for all clinical performance areas are awarded in this same general way. With respect to practice organization, points are often awarded for the presence or absence of some desirable feature; for example, a practice can receive 1.5 points if there are "clearly defined arrangements for backing up computer data" (Roland, 2004, p. 1451). After all points have been determined and summed, the result is multiplied by a predetermined "per point" amount to calculate the total payment to the practice.

We found no studies in the literature that investigated whether paying for performance based on an aggregate measure, or rewarding individual measures, leads to better performance.

Determining the Denominator for Constucting Measures

Which patients should be included when constructing performance measures to use in awarding P4P payments? If patients are included inappropriately, the provider has a financial incentive to give care that may not be needed, which could increase costs and raise issues of patient safety. For example, a common process is to use all patients meeting specific diagnoses criteria (e.g. to establish that the patient has diabetes) in calculating performance on a specific measure. But, this may not always make sense; a physician treating a diabetic patient with terminal cancer arguably should not be penalized if the patient does not receive a scheduled foot exam according to guidelines. In designing P4P initiatives, there are essentially three approaches to addressing this problem. First, threshold performance levels may be set at some target less that 100%. For instance, an 80% threshold would not penalize a provider for using her best clinical judgment in not providing guideline-recommended treatments to up to 20% of her patients. The drawback of this approach is that, if the threshold is set too low, achieving a level of performance necessary to receive a P4P reward may be "too easy," and there may be little actual quality improvement as a result. Second, providers can be allowed to formally exclude patients from measurement, if the patients meet predetermined criteria. The drawback of this approach is that the criteria for exclusion may be too general, creating considerable latitude for the provider to construct the panel of patients to be used for measurement purposes. The result could be provider "gaming" of the process, inappropriately excluding patients who would bring down the average performance of the practice. Third, a statistical risk-adjustment technique can be applied to "level the playing field" among providers in the P4P program. If the risk adjustment approach is effective, the performance of providers is compared for an "average" panel of patients. The drawback of using a statistical risk adjustment methodology is that it may not be transparent to providers, and may cause confusion and suspicion.

Also, statistical risk–adjustment methodologies may not remove enough of the performance variation associated with patient characteristics to adequately address the denominator problem.

Determining the appropriate denominator for measurement purposes is an important design decision because it can have a major impact on the money paid out under P4P programs, as well as which providers receive payments. For example, research suggests that excluding as few as three diabetic patients in a primary care physician's practice can have a large impact on average practice performance (Hofer, et. al, 1999.) And, done incorrectly, it can raise questions about the validity of the entire methodology for allocating reward dollars.

Practice Example

Evaluators of the U.K. P4P initiative have explored issues that can arise in determining the denominator for measuring performance. In this program, physicians were able to use various criteria to exclude individual patients from calculations of quality measures, a process called "exception reporting." The following reasons are permitted for exclusion of patients from the measurement denominator used to reward GP performance (Doran, 2008, p. 276):

1. "The patient has received at least three invitations for review during the preceding 12 months but has not attended."

2. "The indicator is judged to be inappropriate for the patient because of particular circumstances, such as terminal illness, extreme frailty, or the presence of a supervening condition that makes the specified treatment clinically inappropriate."

3. "The patient has recently received a diagnosis or has recently registered with the practice."

4. "The patient is taking the maximum tolerated dose of a medication, but the levels remain suboptimal."

5. "The patient has had an allergic or other adverse reaction to a specified medication, but the levels remain suboptimal."

6. "The patient does not agree to the investigation or treatment."

7. "A specified investigative service is unavailable to the family practitioner."

Evidence from the first year of the U.K.'s P4P initiative suggested that "gaming" of the exclusion criteria might have been an issue. The factor that had the greatest impact in explaining variation on performance across practices was exception reporting; an increase in 1 percent in the proportion of patients excluded was associated with an increase of .31% in performance. Overall rates of exception reporting ranged from 0 to 85%, suggesting that at least some practices may have engaged in excessive exception reporting (Doran, et. al., 2006). However, more recent work with better data suggests these fears likely were unfounded. The median percent of patients excluded in the second year of the program was 5.3%, with characteristics of physicians and patients explaining less than 3% of the variation in

exception reporting. The authors estimate that exception reporting "accounted for approximately 1.5% of the cost of the pay-for-performance" program in that year (Doran, et. al., 2008, p. 274).

The U.K. experience with a relatively simple and transparent approach to "leveling the playing field" when measuring performance is promising. Assuring a "level playing field" becomes especially desirable when outcome measures are used in pay-for-performance programs, because then patient behaviors and characteristics are likely to play even more important roles in achieving goals and determining payouts.

Adjusting for Practice Characteristics

It is frequently argued that some practices are better able to respond to P4P programs because they 1) have more resources, financial and otherwise, or 2) their patients are better able to adhere to treatment plans and carry out self-management activities. The latter issue can be addressed, at least conceptually, through risk adjustment techniques that incorporate patient socio-demographic characteristics. However, the absence of a "level playing field" in terms of practice resources is more difficult to address. Potentially, it could have longer-term consequences for the quality of health care received by economically disadvantaged groups of patients.

The concern is that, if disadvantaged patients make up a disproportionately large percentage of patients in some practices, then these practices may receive no payments (if the patients are uninsured) or low payments (if the patients are enrolled in Medicaid) for providing services to many of their patients. These practices then would be less likely to have the financial resources required to make the investments needed to achieve P4P benchmarks and receive P4P awards. If this is the case, the difference in the financial condition of practices serving significant numbers of disadvantaged patients and other practices could widen. That is, P4P could contribute to a situation where the "rich get richer and the poor get poorer." If this occurs, P4P also could contribute to widening existing racial and ethnic disparities in quality of health care.

There are two ways to address this concern. First, as already mentioned, risk-adjustment approaches could be employed, so that awards were based on an "average" patient panel. Second, dollars could be allocated directly to these practices to improve their infrastructures where this seemed warranted.

Practice Example

In the U.K., concern about the ability of practices located in lower income areas to compete effectively for P4P rewards were expressed prior to program implementation. Consequently, researchers in the U.K. looked for evidence of any problems in this regard. In particular, Doran, et al. (2008) examined the relationship between degree of "deprivation" in the census areas in which physician practices were located and performance on quality indicators over the first three years of the program. They found that median achievement levels on the indicators grew for practices in both the lowest and highest income areas and that the gap in achievement between these practices actually narrowed over time. This suggests that, contrary to concerns, the practices in low income areas were

not hurt by the P4P program and their patients benefited from improvement in the quality of care they received. However, as described above, the U.K had an "exception" system in place that may have contributed to this favorable finding.

Also in the U.K., Guilford, et al. (2007) examined the same issue, with a focus on the performance of practices in serving diabetes patients. They found that practices located in "deprived" areas were less successful than other practices in achieving P4P benchmarks in the first year of the P4P programs. This is a different finding than reported by Doran, et. al. (2007) but the two sets of results are not necessarily in conflict; Doran, et. al. (2008) examined data over a longer time period and for more measures.

Srirangalingam, et. al. (2006) conducted an analysis of treatment for diabetes in clinics located in deprived areas of central London, examining how referral patterns changed after implementation of the P4P initiatives. They reported a significant increase in referrals to specialists for patients with poor blood sugar control. It is not clear if this resulted in better quality of care for these patients. Also, the referrals may have changed the panel of patients employed to measure performance in GP practices, improving the probability of practices receiving a P4P award.

There has been much less attention devoted to this issue by program evaluators in the United States, possibly because most "real world" P4P programs have been implemented by health plans and therefore have affected primarily privately insured populations. Several early experimental studies of P4P in the U.S. were carried out in Medicaid environments, but both experimental and control practices served large numbers of Medicaid recipients. Therefore, there was little opportunity to compare the results for practices serving low income populations to other practices. Karve, et. al. (2008) estimated a statistical relationship between a hospital's performance in Medicare's P4P program and the proportion of patients who were African American. They found that having a higher proportion of African American patients was associated with lower levels of performance on P4P indicators related to treatment for AMI and community-acquired pneumonia.

Based on the existing research, the impact of P4P on practices serving disadvantaged populations, as well as its ultimate impact on racial disparities in quality of care and health outcomes, is unclear. This is a question that deserves greater research attention as P4P initiatives expand to encompass all population segments in communities.

REFERENCES

1. Achat H, McIntyre P and Burgress M (1999). 'Health care incentives in immunization'. *Australian and New Zealand Journal of Public Health,* vol 23, pp 285–88.

2. Amundson G, Solberg LI, Reed M, Martini EM and Carlson R (2003). 'Paying for quality improvement: compliance with tobacco cessation guidelines'. *Joint Commission Journal on Quality & Safety*, vol 29, pp 59–65.

3. Armour BS, Friedman C, Pitts M, Wike J, Alley L, and Etchason J (2004). 'The influence of year-end bonuses on colorectal cancer screening'. *The American Journal of Managed Care* vol 10, pp. 617-24.

4. Armour BS, Pitts MM, Maclean R, Cangialose C, Kishel M, Imai H and Etchason J (2001). 'The effect of explicit financial incentives on physician behavior'. *Archives of Internal Medicine*, vol 161, pp 1261–66.

5. Ashworth M, Armstrong D, de Freitas J, Boullier G, Garforth J and Virji A (2005). 'The relationship between income and performance indicators in general practice: a cross-sectional study'. *Health Services Management Research*, vol 18, pp 258–64.

6. Beaulieu ND and Horrigan DR (2005). 'Organisational processes and quality: putting smart money to work for quality improvement'. *Health Services Research,* vol 40, 1318–34.

7. Bhattacharyya T, Mehta P, and Freiberg AA (2008). 'Hospital characteristics associated with success in a pay-for-performance program in orthopaedic surgery'. *The Journal of Bone & Joint Surgery* vol 90, pp. 1240-43.

8. Cameron PA, Kennedy MP and McNeil JJ (1999). 'The effects of bonus payments on emergency service performance in Victoria'. *The Medical Journal of Australia*, vol 171, pp 243–46.

9. Campbell SM, McDonald R, and Lester H (2008). 'The experience of pay for performance in English family practice: a qualitative study'. *Annals of Family Medicine* vol 6, pp. 228-34.

10. Campbell S, Reeves D, Kontopantelis E, Middleton E, Sibbald B and Roland M (2007). 'Quality of primary care in England with the introduction of pay for performance'. *The New England Journal of Medicine*, vol 357, pp 181–90.

11. Casale AS, Paulus RA, Selna MJ, Doll MC, Bothe Jr AE, McKinley KE, Berry SA, Davis DE, Gilfillan RJ, Hamory BH, and Steele Jr GD (2007). '"ProvenCare℠" A provider-driven pay-for-performance program for acute episodic cardiac surgical care'. *Annals of Surgery* vol 246, pp. 613-23.

12. Chiang C-Y, Enarson DA, Yang S-L, Suo J, Lin T-P (2002). 'The impact of national health insurance on the notification of tuberculosis in Taiwan'. *The International Journal of Tuberculosis and Lung Disease*, vol 6, pp 974–79.

13. Christensen DB, Neil N, Fassett WE, Smith DH, Holmes G and Stergachis A (2000). 'Frequency and characteristics of cognitive services provided in response to a financial incentive'. *Journal of the American Pharmaceutical Association,* vol 40, pp 609–17.

14. Christianson JB, Leatherman S, and Sutherland K (2008a). 'Lessons from evaluations of purchaser pay-for-performance programs'. *Medical Care Research and Review*, vol 65, pp 5S-35S.

15. Christianson JB, Leatherman S, and Sutherland K (2008b). *Financial Incentives, Healthcare Providers and Quality Improvements. A Review of the Evidence*. London, England: The Health Foundation.

16. Christianson JB, Leatherman S, and Sutherland K (2007). 'Paying for quality: Understanding and assessing physician pay-for-performance initiatives'. RWJF Research Synthesis Report No. 13. Princeton, NJ: The Robert Wood Johnson Foundation.

17. Chung RS, Chernicoff HO, Nakao KA, Nickel RC and Legorreta AP (2003). 'A quality-driven physician compensation model: Four-year follow-up study'. *Journal for Healthcare Quality*, vol 25, pp 31–37.

18. Collier VU (2007). 'Use of pay for performance in a community hospital private hospitalist group: a preliminary report'. *Transactions of the American Clinical and Climatological Association* vol 118, pp. 263-72.

19. Curtin K, Beckman H, Pankow G, Milillo Y and Greene RA (2006). 'Return on investment in pay for performance: a diabetes case study'. *Journal of Healthcare Management*, vol 51, pp 365–76.

20. Damberg CL, Raube K, Williams T and Shortell SM (2005). 'Paying for performance: implementing a statewide project in California'. *Quality Management in Health Care*, vol 14, pp 66–79.

21. Doran T. and Fullwood C (2007). 'Pay for performance: is it the best way to improve control of hypertension'? *Current Hypertension Reports* vol 9, pp. 360-67.

22. Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U and Roland M (2006). 'Pay-for-performance programmes in family practices in the United Kingdom'. *The New England Journal of Medicine*, vol 335, pp 375–84.

23. Doran T, Fullwood C, Kontopantelis E, and Reeves D (2008). 'Effect of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework'. *Lancet* vol 372, pp. 728-36.

24. Doran T, Fullwood C, Reeves D, Gravelle H, and Roland M (2008). 'Exclusion of patients from pay-for-performance targets by English physicians'. *The New England Journal of Medicine* vol 359, pp. 274-84.

25. Dudley RA, Frolich A, Robinowitz DL, Talavera JA, Broadhead P, Luft HS and McDonald K (2004). *Strategies to Support Quality-Based Purchasing: A review of the evidence.* Technical Review Number 10. AHRQ Publication No. 04-0057, Agency for Healthcare Research and Quality.

26. Ettner SL, Thompson TJ, Stevens MR, Mangione CM, Kim C, Steers WN, Goewey J, Brown AF, Chung RS, Vankat Narayan KM and the TRIAD Study Group (2006). 'Are physician reimbursement strategies associated with processes of care and patient satisfaction for patients with diabetes in managed care?' *Health Services Research*, vol 41, pp 1221–41.

27. Fairbrother G, Hanson KL, Friedman S and Butts GC (1999). 'The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates'. *American Journal of Public Health*, vol 89, pp 171–75.

28. Fairbrother G, Siegel MJ, Friedman S, Kory PD and Butts GC (2001). 'Impact of financial incentives on documented immunization rates in the inner city: results of a randomised controlled trial'. *Ambulatory Pediatrics*, vol 1, pp 206–12.

29. Felt-Lisk S, Gimm G and Peterson S (2007). 'Making pay-for-performance work in Medicaid'. *Health Affairs – Web Exclusive,* June 26, w516–w527.

30. Fisher ES (2006). 'Paying for performance: risks and recommendations'. *The New England Journal of Medicine*, vol 355, pp 1845–47.

31. Gagné M and Deci EL (2005). 'Self-determination theory and work motivation'. *Journal of Organisational Behaviour*, vol 26, pp 331–62.

32. Gené-Badia J, Escaramis-Babiano G, Sans-Corrales M, Sampietro-Colom L, Aguado-Menguy F, Cabezas-Peña C and Gallo de Puelles P (2007). 'Impact of economic incentives on quality of professional life and on end-user satisfaction in primary care'. *Health Policy*, vol 80, pp 2–10.

33. Glickman SW, Ou F-S, DeLong ER, Roe MT, Lytle BL, Mulgund J, Rumsfeld JS, Gibler WB, Ohman EM, Schulman KA and Peterson ED (2007). 'Pay for performance, quality of care, and outcomes in acute myocardial infarction'. *Journal of the American Medical Association*, vol 297, pp 2373–80.

34. Grady KE, Lemkau JP, Lee NR and Caddell C (1997). 'Enhancing mammography referral in primary care'. *Preventive Medicine*, vol 26, pp 791–800.

35. Greene RA, Beckman H, Chamberlain J, Partridge G, Miller M, Burden D and Kerr J (2004). 'Increasing adherence to a community-based guideline for acute sinusitis through education, physician profiling and financial incentives'. *The American Journal of Managed Care*, vol 10, pp 670–78.

36. Grossbart SR (2006). 'What's the return? Assessing the effect of 'pay-for-performance' initiatives on the quality of care delivery'. *Medical Care Research & Review*, vol 63 (1 Suppl), pp 29S–48S.

37. Gulliford MC, Ashworth M, Robotham D and Mohiddin A (2007). 'Achievement of metabolic targets for diabetes by English primary care practices under a new system of incentives'. *Diabetic Medicine*, vol 24, pp. 505-11.

38. Hillman AL, Ripley K, Goldfarb N, Nuamah I, Weiner J and Lusk E (1998). 'Physician financial incentives and feedback: failure to increase cancer screening in Medicaid managed care'. *American Journal of Public Health*, vol 88, pp 1699–1701.

39. Hillman AL, Ripley K, Goldfarb N, Weiner J, Nuamah I and Lusk E (1999). 'The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care'. *Pediatrics*, vol 104, pp 931–35.

40. Hofer TG, Hayward RA, Greenfield S, Wagner EH, Kaplan SH and Manning WG (1999). 'The unreliability of individual physician 'report cards' for assessing the costs and quality of care of a chronic disease'. *Journal of the American Medical Association*, vol 281, pp 2098–105.

41. Karve A, Ou F-S, Lytle BL, and Peterson ED (2008). 'Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients.' *American Heart Journal* vol 155, pp. 571-76.

42. Khunti K, Gadsby R, Millett C, Majeed A and Davies M (2007). 'Quality of diabetes care in the U.K.: comparison of published quality-of-care reports with results of the quality and outcomes framework for diabetes'. *Diabetic Medicine*, vol 24, pp. 1436-41.

43. Kouides RW, Bennett NM, Lewis B, Cappuccio JD, Barker WH and LaForce FM (1998). 'Performance-based physician reimbursement and influenza immunization rates in the elderly. The primary-care physicians of Monroe County'. *American Journal of Preventive Medicine*, vol 14, pp 89–95.

44. Kraft AD, Capuno JJ, Quimbo SA, and Tan Jr CAR (2008). 'Information, incentives and practice patterns: the case of TB dots services and private physicians in the Philippines'. *The Singapore Economic Review* vol 53, pp. 43-56.

45. Langham S, Gillam S and Thorogood M (1995). 'The carrot, the stick and the general practitioner: how have changes in financial incentives affected health promotion activity in general practice?' *British Journal of General Practice*, vol 45, pp 665–68.

46. Larsen DL, Cannon W and Towner S (2003). 'Longitudinal assessment of a diabetes care management system in an integrated health network'. *Journal of Managed Care Pharmacy*, vol 9, pp 552–58.

47. Levin-Scherz J, DeVita N and Timbie J (2006). 'Impact of pay-for-performance contracts and network registry on diabetes and asthma HEDIS measures in an integrated delivery network'. *Medical Care Research & Review*, vol 63(1 Suppl), pp 14S–28S.

48. Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A and Bratzler DW (2007). 'Public reporting and pay for performance in hospital quality improvement'. *The New England Journal of Medicine*, vol 356, pp 486–96.

49. McDonald R, Harrison S, Checkland K, Campbell SM and Roland M (2007). 'Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study'. *British Medical Journal*, vol 334, pp 1357–62.

50. McDonald R, Harrison S, and Checkland K (2008). 'Incentives and control in primary health care: findings from English pay-for-performance case studies'. *Journal of Health Organization and Management* vol 22, pp. 48-62.

51. Mehrotra A, Pearson SD, Coltin KL, Kleinman KP, Singer JA, Rabson B, and Schneider EC (2007). 'The response of physician groups to P4P incentives'. *The American Journal of Managed Care* vol 13, pp. 249-55.

52. Millett C, Gray J, Saxena S, Netuveli G, and Majeed A (2007). 'Impact of a pay-for-performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes'. *Canadian Medical Association Journal* vol 176, pp. 1705-10.

53. Morrow RW, Gooding AD and Clark C (1995). 'Improving physicians' preventive health care behaviour through peer review and financial incentives'. *Archives of Family Medicine*, vol 4, pp 165–69.

54. Nahra TA, Reiter KL, Hirth RA, Shermer JE, Wheeler JRC (2006). 'Cost-effectiveness of hospital pay-for-performance incentives'. *Medical Care Research and Review*, vol 63, pp 49S–72S.

55. Nalli GA, Scanlon DP and Libby D (2007). 'Developing a performance-based incentive program for hospitals: a case study from Maine'. *Health Affairs*, vol 26, pp 817–24.

56. Nelson AR (2007). 'Pay-for-performance programs: ethical questions and unintended consequences'. *Current Clinical Practice*, vol 1, pp 16-8.

57. Parke II DW (2007). 'Impact of a pay-for-performance intervention: financial analysis of a pilot program implementation and implications for ophthalmology (an American Ophthalmological Society thesis)'. *Transactions of the American Ophthalmological Society* vol 105, pp. 448-60.

58. Petersen LA, Woodard LD, Urech T, Daw C and Sookanan S (2006). 'Does pay-for-performance improve the quality of health care?' *Annals of Internal Medicine*, vol 145, pp 265–72.

59. Ritchie LD, Bisset AF, Russell D, Leslie V and Thomson I (1992). 'Primary and preschool immunization in Grampian: progress and the 1990 contract'. *British Medical Journal*, vol 34, pp 816–19.

60. Roland M (2004). 'Linking physicians' pay to the quality of care: a major experiment in the United Kingdom'. *The New England Journal of Medicine*, vol 251, pp 1448–54.

61. Rosenthal MB and Frank RG (2006). 'What is the empirical basis for paying for quality in health care?' *Medical Care Research & Review*, vol 63, pp 135–57.

62. Rosenthal MB, Frank RG, Li Z and Epstein AM (2005). 'Early experience with pay-for-performance: from concept to practice'. *Journal of the American Medical Association*, vol 294, pp 1788–93.

63. Scott A and Hall J (1995). 'Evaluating the effects of GP remuneration: problems and prospects'. *Health Policy*, vol 21, pp 183–95.

64. Simpson CR, Hannaford PC, Lefevre K and Williams D (2006). 'Effect of the U.K. incentive-based contract on the management of patients with stroke in primary care'. *Stroke*, vol 37, pp 2354–60.

65. Srirangalingam U, Sahathevan SK, Lasker SS and Chowdhury TA (2006). 'Changing pattern of referral to a diabetes clinic following implementation of the new UK GP contract'. *British Journal of General Practice,* vol 56, pp 624–26.

66. St. Jacques PJ, Patel N and Higgins MS (2004). 'Improving anesthesiologist performance through profiling and incentives'. *Journal of Clinical Anesthesia,* vol 16, pp 523–28.

67. Sutton M and McLean G (2006). 'Determinants of primary medical care quality measured under the new UK contract: cross sectional study'. *British Medical Journal*, vol 332, pp 389–90.

68. Town R, Kane R, Johnson P and ButlerM (2005). 'Economic incentives and physicians' delivery of preventive care: a systematic review'. *American Journal of Preventive Medicine*, vol 28, pp 234–40.

69. Whalley D, Gravelle H, and Sibbald B (2008). 'Effect of the new contract on GPs' working lives and perceptions of quality of care: a longitudinal survey'. *British Journal of General Practice* vol 58, pp. 8-14.

70. Young GJ, Meterko M, Beckman H, Baker E, White B, Sautter KM, Greene R, Curtin K, Bokhour BG, Berlowitz D, and Burgess Jr JF (2007). 'Effects of paying physicians based on their relative performance for quality'. *Journal of General Internal Medicine* vol 22, pp. 872-76.

Summaries of Published Studies[*]

Achat, McIntyre and Burgress (1999) reviewed use of incentives to influence immunisation uptake, identified issues in developing incentive programmes and examined findings in the context of a new immunisation incentive scheme in Australia. They conducted a MEDLINE search, in English, under immunisation and financial incentives, from 1966 to 1998. They discussed a U.S. study in New York by Kouides *et al* (1998) which found that, when primary care physicians were rewarded for reaching a 70 per cent target with a fee increase of 10 per cent, the average rate was 73.1 per cent compared with 55.7 per cent in a comparison group. Incentives were less influential in practices with fewer than 100 patients. Ritchie *et al* (1992) looked at changes in rates after the implementation of a new contract for GPs in Scotland in 1990. GPs received additional payments of £1800 (high target – 90 per cent) and £600 (low target). The number of physicians achieving 95 per cent or more rose from 31 to 81 per cent for primary immunisation and from 23 to 64 per cent for preschool boosters. The reasons for the increase were not clear, and there were other factors at work in addition to the financial incentives. Based on the discussion provided, it is not possible to determine the strength of the study designs used by these authors.

Armour *et al* (2001) reviewed the impact of explicit financial incentives at the physician-level on resource use (hospital and visits) and quality measures. The literature review was conducted following the Cochrane Collaborative handbook. The authors did not state how many articles were identified through their review, but they discussed two articles related to resource use and four related to quality of care. One article related to resource use was based on data from a survey of medical directors (Hillman, 1989). The second examined the impact of bonus payments at the physician versus the physician group level. Incentives directed at the individual physician-level were found to be the most effective. The authors reported mixed results regarding the four studies where quality measures were used as outcome variables. One study found no impact while another reported that quality of care, measured by children's immunisation rates, improved. The authors noted the very limited amount of research related to the impact of imposing direct financial incentives on physicians.

Dudley *et al* (2004) conducted a literature review of the evidence on strategies to support quality-based purchasing, which includes a review of the literature on use of financial incentives for providers to improve quality. The authors concluded that a performance-based provider payment could 'plausibly be introduced by a purchaser'. A variety of different outcomes were measured across the studies that were reviewed. The authors interrogated MEDLINE and Cochrane databases, as well as databases documenting ongoing work. Eight randomised trials in which the trial used a performance-based payment as the intervention were identified and included in the review. In four of the articles the

---

recipient of the incentive payment was the individual provider, while in the other four the recipient was either a provider or provider group. In the four studies where the incentive targeted the individual provider, there were five positive and two negative results. In the remaining studies there were one positive and two negative results, where 'positive' indicates a result in the desired direction, and 'negative' means there was no significant effect. In seven studies the target for the incentive was the physician. In these studies there were five significant positive effects and four cases of no significant effect. Positive effects were more likely to be observed when the incentive took the form of an addition to fee-for-service payment than when the incentive was paid as a bonus. Seven studies (and nine dependent variables) addressed preventive care.

Petersen *et al* (2006) reviewed the literature on studies where there was an explicit financial incentive to improve quality. They conducted a PubMed search of the English language literature from 1 January 1980 to 14 November 2005. The 17 empirical studies identified were classified according to the level of incentive (for example, physician, group, payment system) and the type of quality measure rewarded. Thirteen of the 17 studies examined process of care measures, most related to preventive care. Five of the six studies of physician-level incentives, and seven of the nine studies of provider group incentives, found partial or positive effects on quality. One study found a negative effect on care for the sickest patients. Results in four studies suggested unintended side effects. No studies examined optimal duration of incentives or their sustained impact after termination. Overall, the authors observed that there were few empirical studies of the effects of explicit financial incentives on quality.

Rosenthal and Frank (2006) reviewed the literature on paying for quality in healthcare, including brief reviews of the pay-for-performance literature in other fields as well. In 2003, the authors examined the peer-reviewed empirical literature using five databases: MEDLINE, EconLit, ABI Inform, PsychInfo and the Social Science Citation Index. Additional citations were found by examining the reference lists of articles. The review focused on studies that assessed quality-based payment schemes. Studies were excluded that assessed the impact of payment systems on quality of care. The authors located seven published, peer-reviewed empirical studies of the effects of paying for quality in healthcare. Another study located by the review related to contracting for substance abuse treatment, but the rewards were not spelled out so it was excluded from the review. The authors concluded that the empirical foundations for pay-for-performance in healthcare are 'rather weak'. There were only two positive findings and studies with the strongest research designs were more likely to find no impact related to financial incentives. However, the studies were narrowly focused and tended to relate to preventive care. Their implications for more recent pay-for-performance initiatives are not clear.

Scott and Hall (1995) reviewed the effects of different payment methods on GPs using a variety of measures of costs and outcomes of care. Four sources were used in searching the literature: MEDLINE, Social Sciences Citations Index, citations in articles received and citations known to authors. Studies were identified that examined actual changes in GP reimbursements or differences in GP reimbursed in different ways. The authors did not summarise their findings across these studies. Their main conclusion was that, based on the literature, it was not possible to make recommendations about optimal payment systems. Much more research is needed. According to the authors the most 'fundamental criticism' of the literature was that it didn't say whether patients were better or worse as a result of reimbursement

changes. Only one study attempted this, comparing actual practice with clinical guidelines. This study, by the Department of Health in the U.K. (1991), found that GPs were more likely to hit some payment targets when paid specifically to do so. However, this was a before–after study with no controls and unclear data sources.

Town *et al* **(2005)** reviewed the impact of financial incentives on preventive care delivery. A unique aspect of the review is that it is limited to randomised trials. There were eight different financial interventions identified in the review. The incentives included direct payments or bonuses to providers, as well as more diffuse incentives. The authors searched EconLit, Business Source Premier, PsychInfo and MEDLINE. Reference lists were reviewed to identify other articles. The search focused on English language articles published from 1966 to 2002 that addressed primary or secondary prevention or health promotion. Studies using interventions with multiple components, where it was not possible to identify the effect of financial incentives, were also excluded, as were studies that compared outcomes under different payment systems. Two independent reviewers abstracted each article. Only six studies met the inclusion criteria and they generated eight different findings. Of the eight different financial incentives reviewed, only one led to a significantly greater provision of services. The authors noted that this doesn't necessarily imply that financial incentives won't motivate physicians to provide more preventive care. The incentives in the study were weak as the rewards were small. They concluded that small rewards probably won't motivate doctors to change their practices with respect to preventive care.

SUMMARIES OF EXPERIMENTAL STUDIES OF PAY-FOR-PERFORMANCE

Fairbrother *et al* **(1999)** examined the effect of different financial incentives on immunisation coverage, specifically the percentage of children up-to-date on a variety of immunisations. Physicians were assigned to one of three groups: bonus and feedback, enhanced fee-for-service and feedback, and feedback only. Physicians were randomly assigned to the three groups and immunisations were measured at three points in time, approximately four months apart. Nine neighbourhoods in New York City with the highest poverty rates were selected as study sites. Eighty-three physicians were invited to join the study and 61 accepted. Data were collected through chart review. Logistic and linear regression models were used to evaluate outcomes. There was a 25 per cent improvement in up-to-date immunisations in five categories for the bonus group, with no significant changes in the other groups. Much of the improvement appeared to be the result of better documentation.

Fairbrother *et al* **(2001)** conducted a follow-up to a previous study to analyse whether bonus payments and enhanced fee-for-service improved immunisation rates for children, specifically the percentage of children with up-to-date coverage on immunisations. Bonus payments were $1000 and $2500 for 30 point and 45 point improvements, $5000 for reaching 80 per cent up-to-date coverage and $7500 for reaching 90 per cent up-to-date coverage. In the enhanced fee-for-service group, physicians received $5 for each vaccine administered within 30 days of its due date and $15 for each visit at which all due vaccines were administered. A control group received feedback. A previous study by the authors left questions unanswered including: would the improvements in a bonus group continue, would actual practices (as opposed to documentation) improve over time and would the enhanced fee-for-service

group begin to have an impact? Also, the previous study focused on inner city children served under Medicaid. This study included all payers. Incentives were given to 57 randomly selected physicians in New York City four times at four-month intervals based on performance for 50 randomly selected children in their practices. Logistic regression models and linear regression models were used to analyse the data. The lower response rate in this study (compared with the authors' previous study) was a limitation. Both types of financial incentives increased documented immunisations. The authors concluded that the incentives were not sufficient to overcome entrenched physician behaviour patterns and that true immunisation coverage was higher than documented in charts.

Grady *et al* **(1997)** evaluated the success of three different approaches designed to increase referrals by primary care physicians of patients 50 years and older for mammograms. The experiment added what the authors call a token reward for referrals to a strategy of 'cue enhancement' and education. The reward was a cheque based on the percentage referred in a given audit period (for example, $50 for a 50 per cent referral rate). In order to have a comparison group, the rewards for some were not initiated until the second year of the experiment. The study was based on a randomised trial involving 61 practices in Dayton, Ohio and Springfield, Massachusetts over a three-year period resulting in a sample of 11,426 patients. The actual years covered in the analysis are not mentioned. Various statistical techniques were employed, primarily repeated measures ANOVA (analysis of variance), to test for statistically significant differences among the groups. Chart stickers were effective in increasing referrals. Peer-performance and feedback, combined with a reward, did not increase referrals over cueing alone. The authors speculated that the reward offered may have been too small and isolated to have had an impact.

Hillman *et al* **(1998)** evaluated a randomised controlled trial of an intervention intended to improve compliance with four preventive care screening exams for women 50 years and older, with financial incentives for physicians being part of the intervention. The three intervention sites with the highest compliance scores received a full bonus of 20 per cent of capitation. The three with the next highest scores and the three that improved the most got a 10 per cent bonus. Bonuses ranged from $570 to $1260 a site with an average of $775 per audit. Seventeen of the 26 sites received a bonus. Half the 52 primary care sites received the intervention, which included written feedback along with the financial bonus. The study was conducted from 1993 to 1995 in Philadelphia. Tests for the significance of differences in group means were carried out. Financial incentives and feedback did not improve physician compliance. The magnitude of the incentive may have been too small or the physicians may not have been aware of the change in incentives. Both groups saw dramatic increases in preventive care during the study period reflecting national initiatives.

Hillman *et al* **(1999)** conducted a randomised trial of two different interventions, one of which involved a financial incentive, to improve paediatric preventive care in a Medicaid population. The three practice sites with the highest total compliance scores with recommended practices received the full bonus of 20 per cent of capitation. The next three received a 10 per cent bonus, as did the three sites showing the greatest improvement, provided scores increased by at least 10 per cent. Bonuses ranged from $772 to $4682, with an average of $2000. Thirteen of 19 sites received at least one bonus and six sites received two bonuses. The purpose of the study was to determine if a system of semi-annual assessment and

feedback, coupled with financial incentives, could improve paediatric preventive care guidelines as evaluated by semi-annual chart audits from 1993 to 1995. Fifty-three primary care sites in Philadelphia were assigned to three groups: feedback plus incentive, feedback only and a control group. Chart audits were performed at six-month intervals for 18 months. A statistical comparison of means was carried out. Neither intervention resulted in improved care. The authors noted that only 56 per cent of sites reported awareness of the programme despite repeated mailings.

Kouides *et al* **(1998)** conducted an empirical evaluation of the impact of a 1990 Medicare influenza project set in Rochester (New York State), with randomisation of physicians to a control group and an incentive group. Physicians could receive an additional $0.80 per shot or $1.60 per shot if practice immunisation rates of 70 per cent and 85 per cent were achieved respectively. The study took place in 1990 and 1991. Multiple regression techniques were employed in the analysis of physician reports of immunisations. A survey of physician offices was conducted to gather data on practice characteristics. The mean immunisation rate for practices in the incentive group was 68.6 per cent compared with 62.7 per cent for the control group. The median improvement was 10.3 per cent in the incentive group and 3.5 per cent in the control group. The authors conclude that, although the financial incentive was modest, it improved immunisation rates by about 7 per cent.

SUMMARIES OF EVALUATIONS OF P4P INITIATIVES OF HEALTH PLANS AND OTHER PAYERS

Physicians

Amundson *et al* (2003) analysed a programme to reward physicians in a single HMO to advise smokers to quit. Physician groups received bonus payments for achieving target scores on various quality indicators, including providing advice to patients to quit smoking. The authors did not indicate the amounts received by groups specifically for improving in this area. Audits of 14,489 ambulatory patient records were undertaken in 19–20 medical groups from 1996 to 1999. Statistical tests of before–after group means were conducted. Identification of tobacco use in patient records increased from 49 to 73 per cent, and advice to quit increased from 32 to 53 per cent. The number of medical groups in which 80 per cent of patient targets were met increased from zero to eight. The impact of financial incentives by themselves on provider behaviour was difficult to determine because the strength of the incentive was not clear, the incentive at the individual physician-level (as opposed to group) was unclear and the intervention was multi-faceted.

Armour *et al* (2004) conducted a retrospective claims analysis of the impact of physician eligibility for receipt of a bonus payment and performance of colorectal screening. A year-end bonus program was implemented for physicians in a managed care plan. Physicians received bonus payments for conducting colorectal screening for patients who turned 50 years of age. The exact nature of the bonus, including the amount, was not described in the study. The key study variable was "eligible to receive a bonus" with not all managed care physicians qualifying to be eligible. The health plan treated these criteria as proprietary. A multinomial logistic regression model was used to estimate the impact of physician eligibility for the bonus on patient receipt of colorectal screening controlling for patient and physician practice characteristics. Screening use increased significantly in the year after the bonus program was introduced; a 3 percentage point increase, or a 12.8 per cent relative increase. There was no comparable increase for patients of physicians not eligible for the bonus program.

Ashworth *et al* (2005) conducted a multivariate analysis of the relationship among factors related to physician incomes and achievement of quality performance indicators in an inner city health authority in the U.K.. The income of GPs depended on the number of patients in the practice, staff expenses and payments for performance. The time period was two years before new, higher payments for performance were instituted in the NHS. Data were collected on 151 practices in an inner city health authority for 2001 and 2002. Regression analysis, including path analysis, was used to explore relationships. The authors concluded that GPs were able to maximise their incomes by taking on more patients. Achievement of performance targets had little impact. Higher staff budgets were associated with better performance on quality indicators, suggesting that the rewards for performance, which were not large, were offset by the higher costs of achieving higher quality.

Beaulieu and Horrigan (2005) estimated the impact of a managed care organisation's programme, which combined financial incentives and practice support, on the quality of diabetes care. Physicians who met targets or demonstrated significant improvement received a bonus. The largest payment was equivalent to a 12 per cent increase in their per member per month payment (true for both fee-for-

service and capitated physicians). Actual payments ranged from $3000 to $12,000. Data on patient outcomes were self-reported by physicians, with limited audits of medical charts. The control group data were collected as part of the health plan's HEDIS (Healthcare Effectiveness Data and Information Set) reporting. Analysis consisted of statistical comparison of group means before and after the programme. There were significant improvements on five of six process measures. Thirteen of 21 physicians received a financial award. Of the eight not receiving rewards, six improved their scores. There was no evidence that quality declined in areas of care not being measured. Self-selection of physicians into the pilot programme and the small sample size limits the ability to generalise from the results. It also is impossible to determine the marginal effect of the financial incentive because it was implemented along with other practice supports for diabetes care.

Campbell *et al* (2007) assessed trends in quality of care indicators in physician practices in England before and after introduction of a pay-for-performance programme for GPs in 2004. GPs received payments based on the number of points they garnered in the course of a year. Points were awarded for practice structures supporting quality, process of care measures and access measures. Data from 1998, 2003 and 2005 were collected for 42 primary care practices in England for clinical indicators associated with coronary heart disease, asthma and Type 2 diabetes. Trend analysis was conducted for indicators that were eligible for reward under the pay-for-performance programme and also for some indicators that were not. Several different statistical methods were used to test the robustness of the findings. There was a statistically significant, but modest, increase in the trend rate for asthma and Type 2 diabetes indicators after the introduction of the pay-for-performance programme. The lack of a significant increase for coronary heart disease could be due to the fact that scores on these indicators were high prior to the pay-for-performance programme. There was no difference in the trend rates for indicators subject to pay-for-performance and for those that were not. The authors suggest that their analysis may underestimate the impact of pay-for-performance as practices may have implemented some changes in 2003 in anticipation of programme implementation. The lack of a difference between the trends for the two groups of indicators suggests that increases may not be due to pay-for-performance. Alternatively, practitioner attempts to improve scores on pay-for-performance indicators could have had a beneficial 'spillover' effect on other non-measured components of quality. The authors conclude that their results support the view that pay-for-performance can 'make a useful contribution to improving quality' as 'part of a comprehensive quality improvement program' (p 189).

Campbell, McDonald, and Lester (2008) used interviews to explore physician and nurse beliefs and concerns subsequent to implementation of the U.K.'s pay-for-performance program. The pay-for-performance program in the U.K. rewards general practitioner practices for accumulating points by achieving target levels of performance relating to clinical quality, practice administration and other areas. Practice bonuses are directly related to the number of points practices accumulate. Forty-three health professionals (22 physicians and 21 nurses) in 42 practices were invited to participate in the study. The 42 practices employed 110 physicians and 71 nurses. Interviews were conducted between February and August 2007. The authors found agreement that the incentives had been enough to motivate changes in behavior. There was some resentment on the part of nurses that physicians tended to keep the bonus payments rather than distributing them to practice employees. There was also

concern that trying to achieve points had changed the nature of the physician visit, possibly making it more difficult to address the patient's full agenda. There was some indication of reduced continuity of care and greater care fragmentation.

Casale *et al* **(2007)** analyzed a financial incentive program designed to reward increased quality within a single managed care plan. A fixed price for care of coronary artery bypass surgery was implemented, combined with target performance standards that were part of the "care package". The authors sought to determine if the new payment per episode of care approach, combined with the introduction of prompts in an electronic records system and a patient engagement program, could improve quality of care. Mean values were compared before and after program implementation for statistically significant differences. Care improved from an initial 59 per cent of patients receiving all 40 best practices to 86 per cent after 6 months. There were improvements in 30 day clinical outcomes but only "likelihood of discharge to home" was statistically significant.

Chiang *et al* **(2002)** described changes in reporting of TB in Taiwan from 1995 to 1999. Clinicians and hospitals received NT$250 for each confirmed case of TB reported. The authors plotted the number of reported cases from 1995 to 1999. The payment for reporting began in 1997. Changes between various reporting periods (six months) were calculated. There were no tests of significance reported. The incentive programme appeared to have its intended effect. There was a 47 per cent increase in reporting the year that the programme was instituted. However, respectively in 1998 and 1999 the number of reported cases declined slightly (7 per cent and 3 per cent). The study found an impact that was attributed by the authors to incentives. The result was probably significant, but no tests were performed, nor were any data presented concerning the nature of the increased number of reports.

Christensen *et al* **(2000)** carried out an evaluation of an intervention among pharmacists in the State of Washington that involved a financial incentive for providing cognitive services to Medicaid recipients at the time prescriptions were filled. Compensation was $4 for interventions up to six minutes and $6 for longer consultations. All pharmacies also received $40 per month for documenting the cognitive services they provided. Pharmacies were randomly assigned to a study (110) or control (90) group. Cognitive services documentation was audited for completeness and consistency. There was a significant difference in the number of cognitive services per 100 prescriptions (1.59 versus 0.67) and 75 per cent of consultations were less than six minutes. The authors do not state expectations directly, but imply that more consultative services results in better quality care, especially in improving patient safety.

Chung *et al* **(2003)** conducted a qualitative and quantitative assessment of a physician recognition programme employed in the Hawaii Medical Service Association. Physicians received points for achievement relative to quality indicators, patient satisfaction, business operations and utilisation of services. Physicians were ranked and the average incentive reward ranged from 0 to 5.5 per cent, with an average of 3.5 per cent. In 2001 a bonus of up to $3000 was added for practitioners who improved scores. There were payment caps to avoid higher payments to high-fee specialists. Administrative claims data were used to measure the quality indicators and utilisation. A survey was used to collect data on patient satisfaction. Non-parametric tests of statistical significance were conducted for the years 1998–2001. The programme started in 1997. The authors reported results on a subset of indicators (n=3).

These are common measures but there is no explanation for why they were chosen for use in this particular case. There was improvement in ACE inhibitor use and in haemoglobin A1c testing. The results were mixed regarding improvement for an immunisation measure, a finding the authors attribute to external factors. There was no control group, so it is not clear if the improvements were due to the compensation programme. Only a subset of results is provided.

Collier (2007) compared performance by a hospitalist group before and after a reward system for quality was instituted, with hospital groups not under the same contractual incentives used as comparison groups. The exact nature of the reward system was not described. The author notes that rewards were tied to a variety of different types of performance measures, including access, timely records completion, meeting attendance, and meeting quality standards. Performance levels were tracked before and after the contract was put in place and compared with the performance of 2 groups not subject to financial incentives. After one year the hospitalist group under contract improved in all administrative areas, with no similar improvement in the comparison groups. All groups improved by a similar amount with respect to clinical indicators. Because improvement in these areas requires changes in clinical processes, the author speculates that it may take longer than one year to occur.

Curtin *et al* (2006) analysed the cost savings from a pay-for-performance programme directed at physicians providing diabetes care. Payments from a health plan to an individual practice association (IPA) withheld dollars which were then returned to the IPA if it met target performance levels. Each year about $15m of these withheld funds were distributed to 3700 participating physicians, specialists as well as generalists. An average primary care physician's distribution ranged from $6000 to $18,000 annually across all performance measures. Diabetes care was one component of the overall performance score on which payout was based. Historical trend data (2000–02) were used to estimate what the costs of care would have been for diabetes patients in 2003/04 in the absence of the pay-for-performance programme, and this was compared with the cost of the diabetes programme. Claims data provided by the health plan were used in the analysis. Savings were calculated from the perspective of the health plan. The authors found a positive return on investment of 1.6 to 1 in 2003 and 2.5 to 1 in 2004. The most significant cost reductions occurred in the area of hospital care. The authors pointed out that in most instances the pay-for-performance programme essentially rewarded physicians for providing more care for their patients with diabetes, presumably adding to direct treatment costs. Thus, the positive rate of return was more impressive than if achieving the performance goals had required no additional treatment or reductions in treatment.

Damberg *et al* (2005) presented early descriptive findings of the Integrated Health Association's pay-for-performance initiative, with discussion of design and implementation issues. Health plans used a common set of measures drawn from HEDIS to reward physician groups for performance, with public report cards distributed at the same time. Improvements in measures were reported. Tests of significance were referred to but specific results were not provided. Data were from the first reporting year (2003). There was significant improvement on at least four clinical measures for three quarters of the reporting groups. This article focused more on design and implementation issues than on analysis of improvements in quality measures.

Doran *et al* **(2006)** examined the first year experience of family practice doctors in the U.K. in achieving targets under the NHS's new pay-for-performance scheme. In 2004 the NHS committed about $3.2b in new funding for three years for a pay-for-performance programme for family practice doctors. Physicians were rewarded for their performance on 146 quality indicators relating to clinical care for ten chronic diseases, organisation of care and patient experience. Points were awarded on a sliding scale within a payment range, with payment limited to $133 per point awarded in 2004/05, adjusted for disease prevalence. In that period, the maximum that a GP could receive from the programme was $139,400. Data were extracted from a national computer database. Data for exception reports was imputed. Linear multiple least-squares regressions with robust estimates of error variance were used to estimate relationships. Fixed effects for practice location were used. The median practice achieved 95.5 per cent of the points available, in comparison to an expected 75 per cent. Achievement was higher in practices with a high ratio of family practitioners to patients, but all significant effects were small and only 20 per cent of the variance was explained by the regression models. The factor with the greatest effect was exception reporting. Physicians who excluded a large proportion of patients from the calculations performed better. The programme increased the gross income of physicians by an average of $40,200. There were no baseline data in the U.K. to use in the analysis, but there was evidence that quality was improving prior to the programme.

Doran *et al* **(2008)** used multiple linear regression analysis to examine factors that explain the rate of "exception reporting" in the U.K. pay-for-performance program. The pay-for-performance program in the U.K. rewards primary care physicians based on number of points attained in a given year, with the potential for the program to account for 25 per cent of a physician's annual income. In calculating points, physicians are allowed to exclude certain patients meeting predetermined criteria. This practice is called exception reporting, and there has been concern that physicians might use exception reporting inappropriately to generate higher payments. Average exception reporting was found to be much less that previous studies had suggested. This study was based on 2005-2006 data, while earlier studies were based on data from the first year of the program. Physicians excluded a median of 5.3 per cent of their patients. There was little association between patient and practice characteristics and exception reporting. While the average rate of exception reporting was similar to the first program year, the maximum estimated rate was substantially smaller. In all, exception reporting accounted for about 1.5 per cent of overall program costs. The authors concluded that fears of abuses of the exception reporting process seem founded.

Doran *et al* **(2008)** estimated the relationship between "deprivation" of census area in which a physician practices in the U.K. and change in clinical measures of quality under the U.K. pay-for-performance program. The pay-for-performance program in the U.K. is directed at primary care physicians. Physicians are awarded points for achievement of targets related to clinical and administrative clinical performance, and receive an addition to practice income depending on points received. Achievement was tracked over a three-year period on 48 clinical quality indicators for practices located in different census areas grouped by level of deprivation. Logistic regression was used to calculate the odds of practices being in the highest and lowest quintiles with respect to achievement. Multiple linear regression analysis was used to investigate associations between practice level characteristics and

practice achievement. Median reported achievement levels grew in all practices in the first 3 years of the pay-for-performance program. In year 1, deprivation was associated with lower levels of achievement, but the gap between achievement and deprivation narrowed between the first and third years from 4.0 per cent to 0.8 per cent.

Doran and Fullwood (2007) described performance on measures relating to hypertension in the U.K. pay-for-performance program and discussed new evidence regarding exception reporting in the U.K. program. The pay-for-performance program in the U.K. contains several measures relating to hypertension, including blood pressure targets for patients with hypertension, coronary heart disease, diabetes mellitus, and stroke, along with blood pressure targets for all patients 45 and older. In the program, 173 points (16.5 per cent of the total) relate to blood pressure. A practice achieving all these points would receive a reward of $43,250, or an average of $13,000 per physician. The authors examine the percentages of practices achieving targets and changes in performance from year 1 to year 2. Also, in a secondary analysis, the degree of exception reporting was addressed. Rates of achievement were generally high for blood pressure indicators in year one, with a higher rate of achievement in year two than year one. The poorest performing practices in year one showed the greatest rate of improvement. Exception reporting rates were generally low and there was little evidence of widespread gaming of the reporting system. The authors noted that the targets for blood pressure were "less demanding" than would typically be found in blood pressure guidelines.

Ettner *et al* (2006) estimated the association between reimbursement incentives in 10 managed care plans and process measures for quality of care in diabetes treatment. The incentives faced by physicians were measured by proportion of compensation received from salary, capitation, fee-for-service and performance-based payment. A variety of performance measures were used, including receipt of dilated eye exams, foot exams, influenza immunisations, advice to take aspirin, and assessments of glycaemic control, proteinuria and lipid profile. Data were gathered in 2000 and 2001 through patient, provider groups, and health plan surveys and medical records reviews for 6,194 patients with diabetes. The analyses employed multi-level logistic regression techniques with random intercepts for provider groups and health plans. The most significant analytic problem related to high correlation between the payment variables and organisational type. When organisational type was not controlled for in the analysis, care processes were better when physicians were paid on a salary basis, and when quality/satisfaction scores were used to determine a portion of physician payment. The results were confounded by organisational type. Nevertheless, the authors concluded that, regardless of causality, use of quality/satisfaction scores to determine physician compensation 'may indicate delivery of high quality care for diabetes' (p 1222).

Felt-Lisk *et al* (2007) studied a Medicaid pay-for-performance demonstration involving contracting health plans in California. Providers were rewarded for achieving benchmarks for well-baby visits in the Medicaid population. Four of five plans offering new incentives offered bonuses to contracting entities based on the number of children who met well-baby visit guidelines. The fifth made payments directly to physicians using an existing bonus pool. A difference in difference analysis was used to evaluate impacts where data permitted. Qualitative analysis was used to contrast the approaches taken by the Medicaid plans and the difficulties they encountered. Data covered the period from 2002 to 2005, with

the payment years being 2003 to 2005. There were favourable trends overall in the number of well-baby visits, however the experience of the five plans in the study varied. The more successful programmes had greater rewards for providers and had better communication with providers about the incentive programme. There was little information provided in the article relating to the methodology used to estimate quantitative programme impacts.

Gene-Badia *et al* **(2007)** assessed whether implementation of an incentive scheme to improve quality and aid professional development had an impact on quality of professional life and patient satisfaction. Survey data were collected from 257 primary care teams and their patients in Catalonia, Spain in 2002 and 2003. Multivariate regression techniques were used to analyse the impact of financial incentives on 34 measures of quality of professional life and patient satisfaction with care and care facilities. Perception of support from management increased but so did perception of demands on health professionals. There was little evidence of an effect on patient satisfaction.

Greene *et al* **(2004)** evaluated a health plan programme to increase physician adherence to treatment guidelines for acute sinusitis in an IPA in Rochester (New York State). A scoring system was developed based on 20 per cent patient satisfaction, 40 per cent efficiency and 40 per cent quality measures. From 1999 to 2001 the percentage withhold was 15 per cent. In 2000, the withhold was reduced to 10 per cent for the top 5 per cent of performers and increased to 20 per cent for the bottom 5 per cent. Episodes of care were identified for acute sinusitis among 420,000 HMO patients between 1999 and 2001. Statistical process control charts were used to analyse changes over time, with statistical tests of the magnitude of the observed changes. The 'exception rate' decreased by 20 per cent, with most of the change being a decreased use of ineffective antibiotics. Given the multiple interventions involved, it was not possible to determine the contribution of financial incentives to the change.

Gulliford *et al* **(2007)** described trends in diabetes care in 26 South London practices prior to when the U.K.'s pay for performance program was implemented and in the first year after program implementation, supplemented by a cross-sectional analysis of factors that explained variation in practice performance scores during the first year of the program. Eighteen of the 76 clinical indicators in the program pertain to diabetes. They tracked trends in target achievement, without statistical testing, for the 26 practices. Regression analysis was used to assess factors associated with variation in practice performance on measures of diabetes care for all practices in the U.K. program. Among the 26 practices, there was improvement year to year in blood sugar and cholesterol control, with the largest improvements occurring in the year that the pay-for-performance program was implemented. Practices in deprived areas were less successful in achieving targets in the first year of the pay-for-performance program.

Khunti *et al* **(2007)** compared the findings of a review of quality of care in diabetes treatment before implementation of the U.K.'s pay-for-performance program to outcome measures for diabetes care in the first year of the pay-for-performance program. Diabetes care accounts for 99 of 1,050 possible points under the pay-for-performance scheme. Maximum attainment would constitute less than 2 per cent of overall practice income. Their systematic review of the literature identified six studies of diabetes care before implementation of the pay-for-performance program that met inclusion criteria.

The findings of these studies were displayed in tabular form alongside findings from the first year of the pay-for-performance program. The authors concluded that the incentives in the pay-for-performance program led to substantial improvements in many process of care indicators and all intermediate outcome measures. However, there was no control group and the study design did not permit statistical analysis of trends.

Kraft _et al_ (2008) analyzed physician survey data to determine if the opportunity to gain insurance payments affected the probability that physicians followed a recommended treatment protocol for patients with diabetes (self-reported). Accredited facilities can receive a case-based payment for providing a recommended package of services to patients with tuberculosis. The probability of adopting the treatment protocol for diabetes was estimated using a binary logit model and maximum likelihood techniques. Training was found to be a better predictor of adherence to a treatment protocol for TB when the protocol was a significant departure from past practices, but financial incentives were found to be more effective for practices that have demonstrated clinical competence.

Langham, Gillam and Thorogood (1995) examined changes in the distribution of health promotion financial incentive payments after a programme was implemented in the U.K. in 1993 that focused payments on cardiovascular disease. Payments were associated with the performance of screening and the recording of risk factors. Previously, GPs were paid for holding health promotion clinics. The study examined the distribution of health promotion payments between health services authorities and between general practices. The retrospective study of payments included the periods before and after the change in payment approach. Payments were analysed for 78 practices in one authority and 85 in another. Changes in payments were calculated for two measures of relative need. Statistical comparisons of means were conducted. Health promotion payments were found to be more evenly distributed after the change. Practices in areas with the highest need lost more. In general, the resulting distribution was unrelated to need or treatment given after the change.

Larsen, Cannon and Towner (2003) assessed the impact of a disease management process developed by an integrated health system (Intermountain Healthcare – IHC) on diabetes care. The financial incentive was part of a broad-based care improvement effort that included many components. The incentive was not described in detail, but appeared to be relatively small, representing 0.5 to 1 per cent of physician compensation. About half of that incentive related to diabetes care. The authors reported improvement on the key performance measures that were significant and clinically important. Because of the broad-based nature of the programme, it was not possible to determine the impact of the financial incentives by themselves. However, the incentive was small, and therefore it does not seem likely that it was a major influence on behaviour.

Levin-Scherz, DeVita and Timbie (2006) analysed pay-for-performance contracts with physicians for diabetes and asthma care in Massachusetts. The incentive in the programme was applied at the network level. There was a withhold in provider contracts, often at 10 per cent of fees. In some cases there was the opportunity for bonus payments beyond the fee schedule. Withholds were returned or bonuses

earned depending on performance relative to agreed targets. There was a variety of metrics, but the article focused only on performance on diabetes and asthma care. The analysis used claims data taken from the network's multi-year data warehouse. A difference-in-difference analysis was used for the state comparisons. The years included in the study were 2001 to 2003. There was improvement in the network's diabetes measures relative to the index plan in the state and relative to national plans. There was also improvement in asthma measures, but performance in this area started at a relatively high level. The authors noted problems in using claims data to reward performance and the network was looking forward to having access to patient electronic medical records.

McDonald *et al* (2007) conducted an ethnographic study to assess the impact of the NHS pay-for-performance programme on practice organisation, clinical autonomy and internal motivation of GPs and nurses. Data collection took place in two practices in deprived parts of north-west England. These practices had reputations for high quality care and high scores in the first year of the pay-for-performance programme. Observation was combined with interviews, informal conversations and document review. The authors concluded that implementation of financial incentives did not damage internal motivation of GPs, although nurses expressed more concern. Most GPs did not question the quality targets or their implications for clinical quality.

McDonald, Harrison, and Checkland (2008) conducted case studies of two practices in England following implementation of pay-for-performance. The pay-for-performance program in the U.K. is directed at general practitioners, who can earn points based on achievement relative to an extensive number of performance measures. The bonus payments that practices receive depend on the number of points they accumulate. The authors assessed attitudes and behaviours in general practice, with an emphasis on mechanisms and perceptions of control, using observation, interviews and analysis of documentation. Attitudes towards the pay-for-performance program were found to be generally positive in the two practices, but there was discontent in the practices that employed a stronger surveillance system. Nurses who were given responsibility for achieving targets felt greater pressure.

Mehrotra *et al* (2007) conducted a multivariate analysis of survey data collected from physician practices in Massachusetts. Incentives were provided by health plans in a variety of areas including process of care and utilization. The surveys did not ask about specific program incentives. The multivariate analysis addressed the quality improvement activities undertaken by the physician groups. Survey data were collected through telephone interviews with 79 leaders of physician practices in Massachusetts. The leaders were asked about the types of pay-for-performance incentives they faced and their quality improvement activities. A descriptive analysis of the survey data was provided, as well as a multivariate analysis of the association between incentives and the practices' quality improvement activities. Eighty-nine per cent of practices reported incentives for at least one commercial health plan, typically tied to HEDIS reporting measures. About 2/3 of reported incentives were related to utilization measures. Pay-for-improvement incentives tied to HEDIS measures were positively associated with group quality improvement initiatives.

Millet *et al* (2007) conducted a population-based longitudinal study of the effect of the U.K. pay-for-performance program on delivery of smoking cessation advice and on prevalence of smoking among

diabetic patients. The U.K. pay-for-performance program rewarded general practitioners points for achieving clinical and administrative targets. The size of the bonuses paid to physician practices depended on the number of points achieved. The authors analyzed change in recording of smoking status, in documented smoking advice given by the practice, and in the prevalence of smoking among diabetic patients in a single primary care trust containing 36 primary care practices. Conditional logistic regression analysis was used to analyze the data and adjusted odds ratios were reported. More patients with diabetes had their smoking status recorded, and there was significant improvement in documented smoking cessation advice (48 per cent to 83.5 per cent). The prevalence of smoking decreased significantly from 20 per cent to 16 per cent.

Morrow, Gooding and Clark (1995) studied associations between a multi-faceted intervention and improvements in the preventive healthcare behaviours of physicians in an IPA. The only information provided by the authors regarding financial incentives was that a good score on preventive services increased reimbursement for physicians in the health plan. Chart audits of practices in a four state area were conducted from 1987 to 1990 (the number of practices was not provided). Confidence intervals were calculated. There were improvements in virtually all of the preventive measures. The authors observed that they could not necessarily attribute the improvements to the plan's programmes, including financial incentives, as there were confounding motivations for change in physician behaviours.

Parke (2007) analyzed claims data to assess overall health care costs in an employed group before and after introduction of a pay-for-performance program. Physicians received a higher fee-for-service payment if they treated patients as recommended by an electronic reminder related to evidence- based treatment processes. The amount of the increase was approximately 10 per cent. The authors assessed net changes in fixed and variable expenditures by the employer and employees before and after the program was implemented, comparing mean values over two years, with no statistical analyses. Total expenditures declined even though there were benefit design changes and increased per unit payments for physicians. The contribution of the financial incentives for providers was unclear because many other changes, including patient incentive program, were implemented simultaneously.

Ritchie (1992) tracked immunisations in a single region in Scotland before and after introduction of a new contract for primary care physicians in 1990. In this contract, 'item of service' payments were replaced by target payments to encourage GPs to increase rates of childhood immunisations. The details of the payment change were not described. In the study region, this change was accompanied by a records system that provided feedback to GPs regarding their immunisation performance. Immunisation rates for 95 practices encompassing 313 GPs were calculated for children aged two and five on the first day of each quarter for the seven quarters ending in March 1990 and subsequent three-month periods to September 1991. The analysis was retrospective and descriptive and used data drawn from the computer records maintained by the Grampian region in Scotland. A variety of statistical analyses were conducted using linear, non-linear and logistic regression methods. The practices achieving immunisation rates of at least 95 per cent increased from 31 to 81 per cent for primary immunisations. Achievement of 95 per cent rates for pre-school booster immunisation increased from 23 to 64 per cent. The authors noted evidence of 'sustained improvement' but no change in overall trends. They suggested

that the reasons for the change were complex and should not necessarily be attributed to the new contract and the change in financial incentives introduced by it.

Rosenthal *et al* **(2005)** evaluated the impact of a physician pay-for-performance programme implemented by a health plan. Beginning in July 2003, participants received a quarterly bonus of $0.23 per member per month for each performance target met or exceeded. The overall potential for a group with 10,000 health plan patients was $270,000 per year. This represented about 5 per cent of professional capitation paid by the plan and about 0.8 per cent of the group's overall revenue. The evaluation focused on three process measures of clinical quality: cervical cancer screening, mammography and haemoglobin testing. Within the plan, some medical groups received pay-for-performance payments, while groups in another region did not. Data on performance were available before and after the programme was implemented. Generalised least squares techniques were used to estimate a difference-in-difference model. Compared with the groups not receiving a pay-for-performance payment, the groups receiving payment demonstrated greater improvement only in cervical cancer screening. Because payment was made for achieving benchmarks, groups that improved the least, because they started out at a high level, received the most bonus money.

St. Jacques, Patel, and Higgins (2004) assessed the impact of implementing a programme of physician profiling, reporting and incentives on the behaviour of anaesthesiologists. For each study month physicians were eligible to receive a variable financial payment of $0–500 depending on individual scores relative to one another. The payment was credited to the physician's personal expense account. Performance was tracked in five areas: percentage of first cases of the day at the room before or at start time, percentage of cases where preparation time was less than a target, percentage of cases delayed while waiting for anaesthesiology evaluation, percentage of cases delayed during anaesthesiologist controlled time and percentage of cases delayed while waiting for anaesthesiology attending. Thirty-one anaesthesiologists in a university hospital were tracked for six months. A statistical comparison of means was carried out. Compared to the first month, performance improved on most measures. Because the programme combined profiling with incentives it was not possible to determine the effect of incentives only. The authors did not relate their findings to patient outcomes.

Simpson *et al* **(2006)** analysed of the impact of a new payment scheme for GPs on recording of quality indicators for patients with stroke. The new payment system, introduced in Scotland in 2004, provided payments to practices that developed an accurate register of stroke patients and for the recording of smoking habits, blood pressure and cholesterol levels. There were also payments for reaching blood pressure control targets and other outcomes. Retrospective data from 310 (self-selected) of Scotland's 850 practices were obtained from a central database in 2005, including data for one year before the new incentive system was introduced and one year after. Binary logistic regression was used to calculate odds ratios for recording of data. Documentation increased from 32.3 to 52.1 per cent. There was a large increase among the oldest patients and most affluent patients. Women had larger increases in documentation than men. The authors noted that inequitable recording still persists, with lower recording for women, older patients and more deprived patients.

Srirangalingam *et al* (2006) conducted an empirical analysis of how referral patterns for diabetes care changed after introduction of the new financial reward system in the U.K.. Under the new general medical services contract for primary care in the U.K., primary care physicians receive financial rewards for performance on diabetes-related quality indicators. Referrals from primary care to a hospital-based diabetes service before and after implementation of the new incentive system were tracked. The study setting was a deprived area of London. Referrals were tracked from November 2003 to November 2004. Statistical tests of significant differences at the 0.05 level were carried out. There was no significant impact on the total number of referrals to the specialty clinic, but there was a significant increase in referrals for poor glycaemic control. The authors concluded that the new contract led to an increase in referrals for patients with unacceptable glycaemic control along with a lower threshold for referrals.

Sutton and McLean (2006) assessed factors related to quality scores under a new U.K. primary medical care contract that pays GPs in part based on quality measures using a relatively complicated formula that the authors do not describe. Data were analysed for 60 practices in two NHS areas in Scotland serving a population of 367,000. Linear regression analysis was used to relate quality scores to various characteristics of the population, GP and GP's practice. The most relevant finding is that practices with higher incomes from other sources had lower quality scores. The authors speculate that the incentive effect of the new contract is weaker when income from other sources makes up a larger portion of practice income.

Whalley, Gravelle, and Sibbald (2008) conducted a statistical analysis of a longitudinal survey of physicians in the U.K. to assess changes in attitudes towards a pay-for-performance program. The pay-for-performance program in the U.K. is aimed at primary care physicians, who can receive bonus payments based on achievement of a wide variety of administrative and clinical targets. The study considered changes in physician responses to a survey that asked about job satisfaction, hours worked, opinion of the impact of pay-for-performance on quality, and other items. Ordinary least squares, fixed-effect, panel data, multiple regression models were used to analyze the physician survey data. The authors found improvements in satisfaction with work hours and remuneration. Physicians were more positive about the impact of the contract on quality of care than they had expected to be.

Young *et al* (2007) evaluated the impact of a financial incentive program for primary care physicians on five diabetes performance measures. Each physician had about 5 per cent of fees withheld and transferred to a performance pool. The money was distributed to physicians based on their performance relative to indicators of clinical quality, patient satisfaction, and practice efficiency. The possible return to an internist in 2003 was $5,500-$16,500. Two-way repeated-measures analysis of variance was applied for each performance measure, testing for statistically significant changes in performance levels and trends over three time periods. Comparisons were made with general trends in the state and nationally. The authors concluded that the overall improvements in performance reflected secular trends. A "modest" one time improvement in physician adherence to eye examination recommendations was attributed to the program.

Institutions

Bhattacharyya, Mehta, and Freiberg (2008) conducted a multivariate analysis of hospital characteristics that predict hospital performance in the top 20 per cent of a pay-for-performance program related to hip and knee replacements. The CMS/Premier Hospital Quality Incentive Demonstration awarded hospitals performing in the top 10 per cent nationally with a 2 per cent addition to their DRG based payment. Hospitals in the top 20 per cent, but not the top 10 per cent, received a 1 per cent bonus. Hospitals were graded on three process measures and 3 outcome measures. The dependent variable in the analysis was whether or not a hospital was in the top 20 per cent. Variables for which univariate tests of association were not significant were dropped from further analysis. A logistic regression equation was estimated that incorporated the remaining variables. Hospitals in the top 20 per cent in the pay for performance calculations were more likely to be located in the Midwest and be teaching hospitals. Neither hospital size nor revenues were associated with top performance. Volume of surgeries was associated with being in the top 20 per cent.

Type of Study: Observational

Cameron, Kennedy, and McNeil (1999) analysed the impact of a programme of bonus payments for 21 hospitals for improved provision of emergency services. Beginning in 1995, 21 public emergency departments in Victoria, Australia were given bonus payments at the beginning of each fiscal year. They were required to return portions of the bonus if they were unable to meet targets for emergency care. The payments started at AUD$7.2m in total, and increased to AUD$17m by 1997/98. The targets related to areas of performance such as ambulance bypass, waiting time for patients with different levels of emergency and access block (patients waiting more than 12 hours for admission to a hospital). The authors used regression analysis to examine performance on the set of payment measures for two years before and three years after the bonus programme was initiated. The data were self-reported by the study hospitals and not audited. There was no explanation regarding how the authors specified the regression equations and carried out their statistical tests. The authors found that performance improved in all areas. All the results were significant except for the reduction in access block. These results were sustained over the three-year post-intervention period. The authors attributed the success of the incentive programme in part to the fact that it was developed collaboratively with local providers of emergency care.

Glickman *et al* (2007) analysed whether a hospital pay-for-performance programme implemented by Medicare improved care for patients with AMI. Hospitals in the two highest deciles of performance received a reimbursement bonus while those in the lowest decile risked future financial penalties under Medicare's Hospital Quality Incentive Demonstration, which began in 2003. In the first two years, payments totalling $17.55 were made across five clinical conditions, one of which was AMI. In the first year, 123 hospitals received payments; 115 received them in the second year. Data were used for 500 hospitals already participating in a quality improvement initiative (CRUSADE); 54 of these were in the Medicare pay-for-performance initiative, allowing for the creation of a control group of 446 hospitals. Data covered a period before and after the pay-for-performance demonstration. Each hospital collected data and submitted it to a central database. Six different processes of care measures were evaluated as

the primary outcome measures. The study also included eight measures of care that were not included in the measures Medicare rewarded as part of the demonstration. There were slightly higher rates of improvement for two of the six measures rewarded by Medicare: aspirin at discharge and smoking cessation counselling. There was no significant difference in a composite that included all six measures, nor was there any significant difference in a composite consisting of all measures not rewarded by Medicare. The hospitals in the analysis were all volunteers and were already committed to improving treatment for patients with AMI. The authors concluded that, while there was no evidence of improvement due to pay-for-performance, neither did they find any adverse effects.

Grossbart (2006) evaluated the impact of the CMS (Centers for Medicare and Medicaid Services) demonstration project on performance improvement in hospitals, using hospitals in a single multi-hospital system. Under a three-year demonstration programme instituted in 2003, 278 hospitals were given financial incentives based on 35 quality measures in five clinical areas. For each clinical area, hospitals with composite scores in the top 10 per cent received a 2 per cent bonus payment on top of normal payments. Hospitals in the second decile received a 1 per cent payment. There was a slight downside risk in the third year for hospitals that did not perform above threshold quality scores. The setting was Catholic Healthcare Partners, which has its headquarters in Ohio. Some of its hospitals participated in the pilot, while others did not; hence, these acted as a control group. Analysis was limited to three of the five clinical areas: AMI, heart failure and pneumonia. Performance in the first year (2004) was compared with the previous year using composite scores. The study was based on care provided to 28,925 patients. Data were obtained from the database of the hospital system. A comparison of mean values was conducted. The pace of quality improvement in the pilot hospitals was found to be slightly greater than in the control group.

Karve *et al* (2008) conducted a statistical analysis of the relationship between a hospital's performance in Medicare's P4P program and the proportion of patients who were African American. Medicare provides financial incentives to hospitals whose care performance ranks in the top 20 per cent, and in the top 10 per cent, in specific disease categories. Hospital performance on measures of acute myocardial infarction, community-acquired pneumonia, and heart failure was analyzed for the second quarter of 2004 and the first quarter of 2005. Multivariate logistic regression was used to determine the association between percentage of African American patients in a hospital and the likelihood that the hospital was in the lowest or highest quintile of performance. There was an inverse association found between per cent of African American patients and performance related to acute myocardial infarction and community acquired pneumonia. The authors concluded that the P4P program may be exacerbating existing racial ethnic disparities in hospital care.

Lindenauer *et al* (2007) assessed the impact of a Medicare hospital pay for performance initiative on four composite measures of quality of care. Hospitals performing in the top decile on 33 quality measures relating to five conditions received a 2 per cent bonus payment. Those in the second decile received a 1 per cent bonus and hospitals not performing above the level of hospitals in the lowest two deciles (established in the first year) were penalised from 1 to 2 per cent. Bonuses averaged $71,960 per

year. The set of hospitals in the study included 613 hospitals that voluntarily reported information about quality of care through a national public-reporting initiative; 207 of these also were participants in the Medicare pay-for-performance demonstration. Changes in performance were compared for the two groups of hospitals, using multivariate methods to control for differences in hospital characteristics. After adjustment for hospital characteristics and baseline performance, pay-for-performance was associated with improvements from 2.6 to 4.1 per cent over two years. The main share of bonus payments went to hospitals with the highest performance at baseline, but hospitals at all levels of baseline performance improved. The authors view the improvements as modest and acknowledge that the hospitals volunteered and that their attempt to control for 'volunteer bias' may not have been entirely successful. In analyses that did attempt to control for this possible bias, effects were smaller.

Nahra *et al* **(2006)** estimated the QALYs gained in a patient population hospitalised for heart treatment, relative to the money spent by a health insurer in pay-for-performance payments to hospitals. A variety of assumptions needed to be made to generate the estimates in the paper. Eighty-five hospitals in Michigan received about $22 million in a four-year period. The hospitals were paid for achieving minimum levels of compliance with accepted clinical standards for two heart conditions. The incentive payments were add-ons to the hospital's DRG-related payments from BCBS of Michigan (a national health insurer). The maximum add-on for heart care was 1 to 2 per cent between 2000 and 2003. Thresholds for receipt of payment were increased from year to year to encourage continuous improvement. The authors translated the measured improvements into estimated years of life gained relative to cost of the programme. Data on costs were collected from BCBS. Process measures were collected over a four-year period from 2000 to 2003 based on hospital self-reports. The authors used estimates from the literature to convert the process improvements into estimates of QALYs gained. The cost per QALY was estimated to be between about $13,000 and $30,000, which the authors observed is well under the consensus measure for value of a QALY, indicating that the initiative was cost effective.

Nalli *et al* **(2007)** conducted a descriptive study of a hospital pay-for-performance programme implemented in Maine in 2005. Hospitals received payments from a fund established by the hospitals for reaching an agreed performance level and then bonus payments from employers based on performance against 22 measures encompassing patient satisfaction, patient safety, clinical effectiveness and efficiency. Qualitative data were collected from programme participants and data on distribution of funds were collected from secondary sources. Six of the ten participating hospitals received payments totalling $89,645. The participants believed that the programme led to care improvements in their hospitals. The programme was not continued, but it was expected that health plans would use the measures developed by the programme in their pay-for-performance efforts.

Appendix E: Final Recommendations Report

Inventory of Performance Measures in Current Use for Pay-for-Performance Programs

Note: We would like an opportunity to update this inventory as more information becomes available. In particular, we are seeking unpublished information about measures used by for-profit health plans outside of Minnesota.

1) Introduction

This inventory builds on the quality measures inventory from Task 1. This indentifies the subset of quality measures that are in use for pay-for performance programs that are currently being implemented for physicians and hospitals in Minnesota and for a prominent national example of a community-wide pay-for-performance program, the Integrated Health Care Association (IHA) program in California, one of the oldest and largest examples of a community-wide health plan sponsored pay-for-performance program in US.

<u>Physicians</u>

Physicians in Minnesota are currently or will soon be participating in a number of pay-for-performance programs. The Minnesota Medical Association (MMA) conducts an annual survey of measures used in Minnesota's P4P programs. These programs include:

- Health plan sponsored programs
- Bridges to excellence, supported by MN Community Measurement

In addition, the recently implemented Medicare Physician Quality Reporting Initiative (PQRI) will further expand physician pay-for-performance in Minnesota.

<u>Hospitals</u>

The Medicare Hospital Compare program is the most well established hospital pay-for-performance program in Minnesota. In addition, while in the research and evaluation stage, the state is in the vanguard of development of pay-for-performance initiatives for rural hospitals. There is also at least one example of state health plan sponsored hospital pay-for-performance program.

2) Physician Pay-for-Performance Measures - The Minnesota Medial Association's annual survey of pay-for-performance measures being used in Minnesota is compared first with the measures being used by IHA and next with the Medicare PQRI measures.

<u>IHA measures compared with Minnesota Measures</u>

 The Integrated Healthcare Association (IHA) is a statewide leadership group that promotes quality improvement, accountability, and affordability of health care in California. IHA membership includes major health plans, physician groups, and hospital systems, plus academic, consumer, purchaser, pharmaceutical and technology representatives. The IHA's

principal projects include pay-for-performance, medical technology assessment and purchasing, the measurement and reward of efficiency in health care, and prevention programs directed at obesity[†].

Below is the IHA P4P measurement set approved by the P4P Steering Committee for the 2008/2009 reporting year. The items highlighted in red are those measures that are not currently included on the MMA's P4P matrix:

[†] Reference: http://www.iha.org/index.html.

California's IHA Project	
	Year 6 Measures: **2008 Measurement Year /** **2009 Reporting Year**
Clinical Domain *Measures to be collected,* *reported and recommended* *for payment*	1. Childhood Immunization Status w/ 24-month continuous enrollment 2. Appropriate Treatment for Children with Upper Respiratory Infection 3. Breast Cancer Screening 4. Cervical Cancer Screening 5. Chlamydia Screening in Women 6. Use of Appropriate Medication for People with Asthma 9. Cholesterol Management LDL Screening (includes Pts. w/ Cardiovascular Conditions) 10. Cholesterol Management: LDL Control <100 (includes Pts. w/ Cardiovascular Conditions) 12. Colorectal Cancer Screening 13. Appropriate Testing for Children with Pharyngitis 14. Avoidance of Antibiotic Treatment of Adults with Acute Bronchitis 15. Use of Imaging Studies for Low Back Pain 16. Medication Monitoring (ACE/ARBs, digoxin, diuretics)
Clinical PO Encounter	**3.75** Encounters per member per year (using Encounter Rate by Service Type specs)
Clinical Weighting	40%
Patient Experience *Measures to be collected,* *reported and recommended* *for payment*	1. Getting Appointment with a Specialist 2. Rating of Specialist 3. Timely Care and Service composite 4. Doctor-Patient Interaction composite 5. Care Coordination composite 6. Rating of PCP 7. Rating of all Healthcare 8. Office Staff composite 9. Health Promotion composite
Patient Experience *Weighting*	25% (20% for Improvement; 80% for Attainment)

IT-Enabled Systemness
Measures to be collected, reported and recommended for payment
	1. Data Integration for Population Management
	    a. Reporting Based on Electronic Information
	    b. Identifying Important Conditions
	2. Electronic Clinical Decision Support at the Point of Care
	3. Care Management
	    a. Coordination with Practitioners
	    b. Chronic Care Management
	        c. Continuity of Care
	4. Access and Communication
	    a. Processes
	5. Physician Measurement and Reporting

Weighting	15%
Coordinated Diabetes Care *Measures to be collected, reported and recommended for payment*	1. HbA1c Screening 2. HbA1c Poor Control (>9) 3. HbA1c Control (<8) 4. LDL Screening 5. LDL Control <100 6. Nephropathy Monitoring 7. Diabetes Registry and related activities 8. Diabetes Care Management Program
Weighting	20%
Efficiency Domain	Generic Prescribing
Efficiency Weighting	Separate from quality incentive pool

Reportable Non-Payment Measures
Measures to be Collected and Publicly Reported, but not recommended for Payment
	1. Medicare Measures:
	
	    a. Breast Cancer Screening
	    b. Diabetes Care HbA1c Screening
	    c. Diabetes Care HbA1c Poor Control
	    d. Cholesterol Management LDL Screening *(Includes Medicare Pts. w/ Cardiovascular Conditions*
	    e. Cholesterol Management: LDL Control <100 *(Includes Medicare Pts. w/ Cardiovascular Conditions*
	    f. Nephropathy Monitoring for Diabetic Patients
	    g. Colorectal Cancer Screening

Transition Measures *Measures to be Collected but not Publicly Reported or recommended for Payment. these measures have been tested and approved for addition to the P4P measure set in the following year.*	Clinical: 1. Evidence-Based Cervical Cancer Screening 2. HbA1c Control (<7) Appropriate Resource Use Measures *(will be used to establish a baseline)*: 1. Inpatient Utilization—Acute Care Discharges 2. Inpatient Utilization—Bed Days 3. Outpatient Surgeries Utilization—% Done in ASC 4. Emergency Department Visits 5. Inpatient Readmissions within 30 Days 6. Generic Prescribing

Testing Measures
Measures to be collected for Testing and Analysis
	1. Blood Pressure Control in Patients with Diabetes
	2. Optimal Diabetes Care
	    a. HbA1c < 8
	    b. Blood Pressure <140/90
	    c. LDL <100
	3. Adolescent Immunizations (Tdap, meningococcal)

‡

---

‡ Items in red are not located on the MMA matrix.

Minnesota P4P measures compared with CMS's PQRI measures. (Those measures listed in red represent the measures indicated in the PQRI, but **not** in the MMA matrix[§]).

Medicare Physician Quality Reporting Initiative (PQRI)

The 2006 Tax Relief and Health Care Act (TRHCA) (P.L. 109-432) required the establishment of a physician quality reporting system, including an incentive payment for eligible professionals (EPs) who satisfactorily report data on quality measures for covered services furnished to Medicare beneficiaries during the second half of 2007 (the 2007 reporting period). CMS named this program the Physician Quality Reporting Initiative (PQRI)[**]. The payment is based on increasing the % annual update to the physician's Medicare fee schedule

<div align="center">2009 PQRI Measures List</div>

12-Lead Electrocardiogram (ECG) Performed for Non-Traumatic Chest Pain

12-Lead Electrocardiogram (ECG) Performed for Syncope

Acute Otitis Externa (AOE): Pain Assessment

Acute Otitis Externa (AOE): Systemic Antimicrobial Therapy – Avoidance of Inappropriate Use

Acute Otitis Externa (AOE): Topical Therapy

Advance Care Plan

Age-Related Macular Degeneration (AMD): Counseling on Antioxidant Supplement

Age-Related Macular Degeneration (AMD): Dilated Macular Examination

AMA-PCPI

Antibiotic Treatment for Adults with Acute Bronchitis: Avoidance of Inappropriate Use

Appropriate Testing for Children with Pharyngitis

Aspirin at Arrival for Acute Myocardial Infarction (AMI)

Asthma: Asthma Assessment

Asthma: Pharmacologic Therapy

Back Pain: Advice Against Bed Rest

Back Pain: Advice for Normal Activities

Back Pain: Initial Visit

Back Pain: Physical Exam

Breast Cancer Resection Pathology Reporting: pT Category (Primary Tumor) and pN Category (Regional Lymph Nodes) with Histologic Grade

Breast Cancer: Hormonal Therapy for Stage IC-IIIC Estrogen Receptor/Progesterone Receptor (ER/PR) Positive Breast Cancer

Cataracts: Comprehensive Preoperative Assessment for Cataract Surgery with Intraocular Lens (IOL) Placement

Chronic Kidney Disease (CKD): Blood Pressure Management

Chronic Kidney Disease (CKD): Influenza Immunization

Chronic Kidney Disease (CKD): Laboratory Testing (Calcium, Phosphorus, Intact Parathyroid Hormone (iPTH) and Lipid Profile)

[§] This comparison is a quick reference and has not been tested for detailed matching of numerators/denominators.

[**] Reference: http://www.cms.hhs.gov/PQRI/01_Overview.asp#TopOfPage

Chronic Kidney Disease (CKD): Plan of Care – Elevated Hemoglobin for Patients Receiving Erythropoiesis-Stimulating Agents (ESA)

Chronic Kidney Disease (CKD): Referral for Arteriovenous (AV) Fistula

Chronic Lymphocytic Leukemia (CLL): Baseline Flow Cytometry

Chronic Obstructive Pulmonary Disease (COPD): Bronchodilator Therapy

Chronic Obstructive Pulmonary Disease (COPD): Spirometry Evaluation

Colon Cancer: Chemotherapy for Stage III Colon Cancer Patients

Colorectal Cancer Resection Pathology Reporting: pT Category (Primary Tumor) and pN Category (Regional Lymph Nodes) with Histologic Grade

Community-Acquired Pneumonia (CAP): Assessment of Mental Status

Community-Acquired Pneumonia (CAP): Assessment of Oxygen Saturation

Community-Acquired Pneumonia (CAP): Empiric Antibiotic

Community-Acquired Pneumonia (CAP): Vital Signs

Coronary Artery Bypass Graft (CABG): Antiplatelet Medications at Discharge

Coronary Artery Bypass Graft (CABG): Beta-Blockers Administered at Discharge

Coronary Artery Bypass Graft (CABG): Deep Sternal Wound Infection Rate

Coronary Artery Bypass Graft (CABG): Lipid Management and Counseling

Coronary Artery Bypass Graft (CABG): Postoperative Renal Insufficiency

Coronary Artery Bypass Graft (CABG): Preoperative Beta-Blocker in Patients with Isolated CABG Surgery

Coronary Artery Bypass Graft (CABG): Prolonged Intubation (Ventilation)

Coronary Artery Bypass Graft (CABG): Stroke/Cerebrovascular Accident (CVA)

Coronary Artery Bypass Graft (CABG): Surgical Re-exploration

Coronary Artery Bypass Graft (CABG): Use of Internal Mammary Artery (IMA) in Patients with Isolated CABG Surgery

Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Patients with CAD and Diabetes and/or Left Ventricular Systolic Dysfunction (LSVD)

Coronary Artery Disease (CAD): Beta-Blocker Therapy for CAD Patients with Prior Myocardial Infarction (MI)

Coronary Artery Disease (CAD): Lipid Profile in Patients with CAD

Coronary Artery Disease (CAD): Oral Antiplatelet Therapy Prescribed for Patients with CAD

Diabetes Mellitus: Diabetic Foot and Ankle Care, Peripheral Neuropathy – Neurological Evaluation

Diabetes Mellitus: Diabetic Foot and Ankle Care, Ulcer Prevention – Evaluation of Footwear

Diabetes Mellitus: Dilated Eye Exam in Diabetic Patient

Diabetes Mellitus: Foot Exam

Diabetes Mellitus: Hemoglobin A1c Poor Control in Diabetes Mellitus

Diabetes Mellitus: High Blood Pressure Control in Diabetes Mellitus

Diabetes Mellitus: Low Density Lipoprotein (LDL-C) Control in Diabetes Mellitus

Diabetes Mellitus: Urine Screening for Microalbumin or Medical Attention for Nephropathy in Diabetic Patients

Diabetic Retinopathy: Communication with the Physician Managing On-going Diabetes Care

Diabetic Retinopathy: Documentation of Presence or Absence of Macular Edema and Level of Severity of Retinopathy

Documentation and Verification of Current Medications in the Medical Record

Elder Maltreatment Screen and Follow-Up Plan

End Stage Renal Disease (ESRD): Influenza Immunization in Patients with ESRD

End Stage Renal Disease (ESRD): Plan of Care for Inadequate Hemodialysis in ESRD Patients

End Stage Renal Disease (ESRD): Plan of Care for Inadequate Peritoneal Dialysis

Endarterectomy: Use of Patch During Conventional Endarterectomy

Falls: Plan of Care

Falls: Risk Assessment

Functional Outcome Assessment in Chiropractic Care

Health Information Technology (HIT): Adoption/Use of Electronic Health Records (EHR)

Heart Failure: Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD)

Heart Failure: Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)

Hemodialysis Vascular Access Decision-Making by Surgeon to Maximize Placement of Autogenous Arterial Venous (AV) Fistula

Hepatitis C: Antiviral Treatment Prescribed

Hepatitis C: Counseling Regarding Risk of Alcohol Consumption

Hepatitis C: Counseling Regarding Use of Contraception Prior to Antiviral Therapy

Hepatitis C: HCV Genotype Testing Prior to Treatment

Hepatitis C: HCV Ribonucleic Acid (RNA) Testing at Week 12 of Treatment

Hepatitis C: Hepatitis A Vaccination in Patients with HCV

Hepatitis C: Hepatitis B Vaccination in Patients with HCV

Hepatitis C: Ribonucleic Acid (RNA) Testing Before Initiating Treatment

Hepatitis C: Testing for Chronic Hepatitis C – Confirmation of Hepatitis C Viremia

HIV/AIDS: Adolescent and Adult Patients with HIV/AIDS Who Are Prescribed Potent Antiretroviral Therapy

HIV/AIDS: CD4+ Cell Count or CD4+ Percentage

HIV/AIDS: HIV RNA Control After Six Months of Potent Antiretroviral Therapy

HIV/AIDS: Pneumocystis Jiroveci Pneumonia (PCP) Prophylaxis

Major Depressive Disorder (MDD): Antidepressant Medication During Acute Phase for Patients with MDD

Major Depressive Disorder (MDD): Diagnostic Evaluation

Major Depressive Disorder (MDD): Suicide Risk Assessment

Medication Reconciliation: Reconciliation After Discharge from an Inpatient Facility

Melanoma: Continuity of Care – Recall System

Melanoma: Coordination of Care

Melanoma: Follow-Up Aspects of Care

Multiple Myeloma: Treatment with Bisphosphonates

Myelodysplastic Syndrome (MDS) and Acute Leukemias: Baseline Cytogenetic Testing Performed on Bone Marrow

Myelodysplastic Syndrome (MDS): Documentation of Iron Stores in Patients Receiving Erythropoietin Therapy

Nuclear Medicine: Correlation with Existing Imaging Studies for All Patients Undergoing Bone Scintigraphy

Oncology: Medical and Radiation – Pain Intensity Quantified

Oncology: Medical and Radiation – Plan of Care for Pain

Oncology: Radiation Dose Limits to Normal Tissues

Osteoarthritis (OA): Assessment for Use of Anti-Inflammatory or Analgesic Over-the-Counter (OTC) Medications

Osteoarthritis (OA): Function and Pain Assessment

Osteoporosis: Communication with the Physician Managing On-going Care Post-Fracture

Osteoporosis: Management Following Fracture

Osteoporosis: Pharmacologic Therapy

Otitis Media with Effusion (OME): Diagnostic Evaluation – Assessment of Tympanic Membrane Mobility

Otitis Media with Effusion (OME): Hearing Testing

Pain Assessment Prior to Initiation of Patient Therapy and Follow-Up

Pediatric End Stage Renal Disease (ESRD): Influenza Immunization

Pediatric End Stage Renal Disease (ESRD): Plan of Care for Inadequate Hemodialysis

Percentage of patients aged 18 years and older receiving a surveillance colonoscopy and a history of colonic polyp(s) in a previous colonoscopy, who had a follow-up interval of 3 or more years since their last colonoscopy documented in the colonoscopy report AMA-PCPI/NCQA

Percentage of patients aged 18 years and older with a diagnosis of venous ulcer who were prescribed compression therapy within the 12-month reporting period AMA-PCPI/NCQA

Perioperative Care: Discontinuation of Prophylactic Antibiotics (Cardiac Procedures)

Perioperative Care: Discontinuation of Prophylactic Antibiotics (Non-Cardiac Procedures)

Perioperative Care: Selection of Prophylactic Antibiotic – First OR Second Generation Cephalosporin

Perioperative Care: Timing of Antibiotic Prophylaxis – Ordering Physician

Perioperative Care: Timing of Prophylactic Antibiotics – Administering Physician

Perioperative Care: Venous Thromboembolism (VTE) Prophylaxis (When Indicated in ALL Patients)

Prevention of Catheter-Related Bloodstream Infections (CRBSI): Central Venous Catheter (CVC) Insertion Protocol

Preventive Care and Screening: Advising Smokers to Quit

Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up

Preventive Care and Screening: Colorectal Cancer Screening

Preventive Care and Screening: Influenza Immunization for Patients ≥ 50 Years Old

Preventive Care and Screening: Inquiry Regarding Tobacco Use

Preventive Care and Screening: Pneumonia Vaccination for Patients 65 Years and Older

Preventive Care and Screening: Screening Mammography

Preventive Care and Screening: Unhealthy Alcohol Use – Screening

Primary Open Angle Glaucoma (POAG): Optic Nerve Evaluation

Primary Open-Angle Glaucoma (POAG): Reduction of Intraocular Pressure (IOP) by 15% OR Documentation of a Plan of Care

Prostate Cancer: Adjuvant Hormonal Therapy for High-Risk Prostate Cancer Patients

Prostate Cancer: Avoidance of Overuse of Bone Scan for Staging Low-Risk Prostate Cancer Patients

Prostate Cancer: Three-Dimensional (3D) Radiotherapy

Radiology: Exposure Time Reported for Procedures Using Fluoroscopy

Radiology: Inappropriate Use of "Probably Benign" Assessment Category in Mammography Screening

Rheumatoid Arthritis (RA): Assessment and Classification of Disease Prognosis

Rheumatoid Arthritis (RA): Disease Modifying Anti-Rheumatic Drug (DMARD) Therapy

Rheumatoid Arthritis (RA): Functional Status Assessment

Rheumatoid Arthritis (RA): Glucocorticoid Management

Rheumatoid Arthritis (RA): Periodic Assessment of Disease Activity

Rheumatoid Arthritis (RA): Tuberculosis Screening

Screening for Clinical Depression and Follow-Up Plan

Screening or Therapy for Osteoporosis for Women Aged 65 Years and Older

Stroke and Stroke Rehabilitation: Anticoagulant Therapy Prescribed for Atrial Fibrillation at Discharge

Stroke and Stroke Rehabilitation: Carotid Imaging Reports

Stroke and Stroke Rehabilitation: Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) Reports

Stroke and Stroke Rehabilitation: Consideration of Rehabilitation Services

Stroke and Stroke Rehabilitation: Deep Vein Thrombosis Prophylaxis (DVT) for Ischemic Stroke or Intracranial Hemorrhage

Stroke and Stroke Rehabilitation: Discharged on Antiplatelet Therapy

Stroke and Stroke Rehabilitation: Screening for Dysphagia

Stroke and Stroke Rehabilitation: Tissue Plasminogen Activator (t-PA) Considered

Thoracic Surgery: Recording of Clinical Stage for Lung Cancer and Esophageal Cancer Resection

Treatment for Children with Upper Respiratory Infection (URI): Avoidance of Inappropriate Use

Urinary Incontinence: Assessment of Presence or Absence of Urinary Incontinence in Women Aged 65 Years and Older

Urinary Incontinence: Characterization of Urinary Incontinence in Women Aged 65 Years and Older

Urinary Incontinence: Plan of Care for Urinary Incontinence in Women Aged 65 Years and Older

P4P Programs

Measure Summary

	BCBS MN 2009 program for 2009 dates of service	Bridges to Excellence 2009 for 2008 dates of service	HealthPartners Partners in Progress 2009 dates of service	HealthPartners Partners in Quality 2009 dates of service	Medica Performance-Based Incentive Program 2008 and 2009 Program DOS	Medica Choice Care Quality Improvement Program 2008 and 2009 Program/DOS	Preferred One 2009 dates of service	Ucare 2009 P4P Program 2009 dates of service	DHS Reporting requirements 2009 dates of service	QCare	MN Community Measurement 2008 dates of service	2009 CMS PQRI
1. Advanced Care Plan												x
2. Asthma - Assessment												x
3. Asthma - Pharmacology											x	x
4. Back Pain – Optimal care for acute low back pain					x							
5. Back Pain – CMS Measures Group												x
6. Back Pain – Initial Visit												x
7. Back Pain – Physical Exam												x
8. Back Pain – Advise For Normal Activity												x
9. Back Pain – Advise Against Bed Rest												x
10. Board Maintenance of Certification	x											
11. Body Mass Index (BMI) assessment - pediatrics			x									
12. Body Mass Index (BMI) and Weight Management Plan	x											x
13. Bronchitis: Avoidance of Antibiotic Treatment in Adults											x	x
14. Bronchitis/URI: Avoidance of Antibiotic treatment in Children	x											
15. Cancer – Breast Cancer Hormonal Therapy												x
16. Cancer – Breast Cancer Pathology Reporting												x
17. Cancer – Chronic Lymphocytic Leukemia Baseline Flow Cytometry												x
18. Cancer – Colon Cancer Chemotherapy for Stage III												x
19. Cancer - Colorectal Pathology												x

[††] This comparison is a quick reference and has not been tested for detailed matching of numerators/denominators.

Measure												
20. Cancer - Melanoma Follow-up Care												x
21. Cancer - Melanoma Continuity of Care												x
22. Cancer - Melanoma Coordination of Care												x
23. Cancer - Multiple Myeloma												x
24. Cancer – Myelodysplastic Syndrome (MDS) and Acute Leukemias Baseline Testing												x
25. Cancer – Myelodysplastic Syndrome (MDS) and Acute Leukemias Iron Stores												x
26. Cancer – Prostate Cancer Avoidance of Overuse of Bone Scan												x
27. Cancer – Prostate Cancer Hormonal Therapy												x
28. Cancer – Prostate Cancer 3D Radiotherapy												x
29. Cancer - Oncology Medical and Radiation Pain Intensity Quantified												x
30. Cancer - Oncology Medical and Radiation Plan for Pain Care												x
31. Cancer - Oncology: Radiation Dose Limits to Normal Tissues												x
32. Cardiovascular – LDL screening						x						
33. Cardiovascular - Acute Myocardial Infarction									x			x
34. Cardiovascular – Non-traumatic Chest pain												x
35. Cardiovascular – CAD Oral Antiplatelets												x
36. Cardiovascular – CAD Beta Blocker Therapy in Patients with Prior MI												x
37. Cardiovascular - CAD ACE or ARB Therapy for Patients with Diabetes or LVSD												x
38. Cardiovascular – CAD Lipid Profile												x
39. Cardiovascular - Congestive Heart Failure									x			
40. Cardiovascular – Heart Failure ACE or ARB Therapy for LVSD												x
41. Cardiovascular – Heart failure – beta blocker for LVSD												x
42. Cardiovascular – Heart failure re-admissions (30d all cause)			x									
43. Cardiovascular – CABG CMS Measures Group												x
44. Cardiovascular – CABG Internal Mammary Artery Use												x
45. Cardiovascular – CABG Preoperative Beta-Blocker in Patients with Isolated CABG Surgery												x

46. Cardiovascular – CABG Prolonged Intubation												x
47. Cardiovascular – CABG Deep Sternal Wound Infection												x
48. Cardiovascular – CABG Stroke or Cerebral Accident												x
49. Cardiovascular – CABG Post Operative Renal Insufficiency												x
50. Cardiovascular – CABG Surgical Re-exploration												x
51. Cardiovascular – CABG Antiplatelet Drugs at Discharge												x
52. Cardiovascular – CABG Beta Blocker at Discharge												x
53. Cardiovascular – CABG Lipid Management and Counseling												x
54. Cardiovascular – MN Community Measurement Optimal Vascular Care	x	x		x	x				x	x	x	
55. Child Developmental Screening									x			
56. Child Mental Health Screening									x			
57. Chiropractic – Functional Outcome												x
58. Chronic Kidney Disease – CMS Measure Group												x
59. Chronic Kidney Disease - Lab Testing												x
60. Chronic Kidney Disease - Blood Pressure Management												x
61. Chronic Kidney Disease - Elevated Hemoglobin Plan of Care												x
62. Chronic Kidney Disease – Influenza Immunization												x
63. Chronic Kidney Disease – Referral for AV Fistula												x
64. Chronic Obstructive Pulmonary Disease (COPD) - Spirometry Evaluation												x
65. Chronic Obstructive Pulmonary Disease (COPD) - Bronchodilator Therapy Evaluation												x
66. Colonoscopy - Endoscopy & Polyp Surveillance Patients with a History of Adenomatous Polyps Avoidance of Inappropriate Use												x
67. Community Acquired Bacterial Pneumonia - Vital Signs												x
68. Community Acquired Bacterial Pneumonia - O2 Saturation												x
69. Community Acquired Bacterial Pneumonia - Mental Status Assessment												x
70. Community Acquired Bacterial Pneumonia - Empiric Antibiotic												x
71. Catheter-Related Bloodstream Infections (CRBSI) - Central Venous Catheter (CVC) Insertion Protocol												x

Measure												
72. Decision Support – Bronchitis avoidance of antibiotic treatment in adults			x									
73. Depression Care – Optimal depression care	x	x		x						x		
74. Depression Care –Diagnostic Evaluation												x
75. Depression Care –Antidepressant During Acute Phase												x
76. Depression Care – Suicide Risk Assessment												x
77. Depression Care –Symptom Assessment Tool												x
78. Diabetes – MN Community Measurement Optimal Care	x	x			x	x	x	x	x	x		
79. Diabetes CMS Measure Group												x
80. Diabetes – Hemoglobin A1c Poor Control												x
81. Diabetes – LDL Control												x
82. Diabetes – High Blood Pressure												x
83. Diabetes – Dilated Eye Exam												x
84. Diabetes – Urine Screen for Microalbumin for Nephropathy												x
85. Diabetes – Foot and Ankle Care Neurological Exam												x
86. Diabetes – Foot and Ankle Care Ulcer Prevention Footwear												x
87. Diabetes – Foot Exam												x
88. E-prescribing – Adoption and Use												x
89. Ear Care – Acute Otitis Externa Topical Therapy												x
90. Ear Care – Acute Otitis Externa Pain Assessment												x
91. Ear Care – Acute Otitis Externa Antibiotic Use												x
92. Ear Care – Acute Otitis Media	x											
93. Ear Care – Otitis Media with Effusion Tympanic Membrane Mobility												x
94. Ear Care – Otitis Media with Effusion Hearing Testing												x
95. Elder Care – Health Evaluation									x			
96. Elder Care – Maltreatment Screen and Follow-up												x
97. Endarterectomy – Use of Patch												x
98. End of life discussion – decision points			x									

174

Measure												
99. End Stage Renal Disease (ESRD) -Influenza Immunization												x
100. End Stage Renal Disease (ESRD) -Inadequate Hemodialysis												x
101. End Stage Renal Disease (ESRD) -Inadequate Peritoneal Dialysis												x
102. End Stage Renal Disease (ESRD) Pediatric – Plan for Inadequate Hemodialysis												x
103. End Stage Renal Disease (ESRD) Pediatric – Influenza Immunization												x
104. Eye Care – Age Related Macular Degeneration Dilated Macular Exam												x
105. Eye Care – Age Related Macular Degeneration Counseling on Antioxidants												x
106. Eye Care – Cataracts Comprehensive Preoperative Assessment												x
107. Eye Care – Diabetic Retinopathy Macular Edema												x
108. Eye Care – Diabetic Retinopathy Communication with Diabetes Physician												x
109. Eye Care – Primary Open Angle Glaucoma Optic Nerve Evaluation												x
110. Eye Care – Primary Open Angle Glaucoma Reduction of Intraocular Pressure												x
111. Falls Risk Assessment												x
112. Falls Care Plan												x
113. Functional Status measurement - PT				x								
114. Generic Drug Use			X	X								
115. Generic Prescribing Provider Decision Support			x									
116. Health information Technology (HIT)	x		x								x	x
117. Hemodialysis – Vascular Access Decision Making												x
118. Hepatitis C –Confirmation												x
119. Hepatitis C – RNA Testing Before Treatment												x
120. Hepatitis C – HCV Genotype Testing Before Treatment												x
121. Hepatitis C – Antiretroviral Treatment Prescribed												x
122. Hepatitis C – HCV Testing within 12 Weeks of Treatment												x
123. Hepatitis C – Counseling Risk of Alcohol Consumption												x
124. Hepatitis C – Counseling Contraception												x

Measure												
125. Hepatitis C – Hepatitis A Vaccination in Patients with HCV												x
126. Hepatitis C – Hepatitis B Vaccination in Patients with HCV												x
127. HIV /AIDS – CD4+ Cell Count or CD4+ Percentage												x
128. HIV / AIDS – Pneumonia Prophylaxis												x
129. HIV / AIDS – Adolescents and Adults Prescribed Potent Antiretroviral												x
130. HIV / AIDS – HIV RNA Control After 6 Months of Potent Antiretroviral												x
131. Hypertension											x	
132. Hysterectomy – non-cancerous hysterectomy post operative complications				x								
133. Immunization Rate – Adolescents										x		
134. Immunization Rate – Children				x				x		x	x	
135. Immunizations- Influenza immunization for Adults												x
136. Immunizations- Pneumonia Vaccination for Adults												x
137. Innovations in Health Care				x								
138. Language and Race Documentation			x									
139. Medication Reconciliation – After Discharge, Inpatient Facility												x
140. Medication Reconciliation – Documentation in Medical Record												x
141. MNCM Direct Data Submission Participation			x									
142. Orthopaedic Care – post operative complications				x								
143. Osteoarthritis – Function and Pain Assessment												x
144. Osteoarthritis – Assess Use of Anti-inflammatory or Analgesic OTC Medications												x
145. Osteoporosis –Communication with Post Fracture Physician												x
146. Osteoporosis –Screening or Therapy												x
147. Osteoporosis – Management Following Fracture												x
148. Osteoporosis - Pharmacologic Therapy												x
149. Pain: Assessment Prior to Initiation of Therapy and Follow-up												x
150. Patient Satisfaction/ Experience				x							x	

151. Patient shared decision making process			x									
152. Perioperative Care – CMS Measures Group												x
153. Perioperative Care – Timing of Antibiotic Prophylaxis, Ordering Physician												x
154. Perioperative Care – Selection of Prophylactic Antibiotic, First OR Second Generation Cephalosporin												x
155. Perioperative Care – Discontinuation of Prophylaxis Antibiotics (non cardiac patients)												x
156. Perioperative Care – VTE Prophylaxis When Indicated in All Patients												x
157. Perioperative Care – Timing of Prophylactic Antibiotics, Administering Physician												x
158. Perioperative Care – Discontinuation of Prophylactic Antibiotics (cardiac patients)												x
159. Pharyngitis – Appropriate Testing for Children											x		x
160. Pneumonia, ED									x				
161. Preventative Care CMS Measures Group													x
162. Preventive services and BMI assessment combo-pediatrics				x									
163. Radiology – Exposure Time Reported for Procedures Using Fluoroscopy													x
164. Radiology – Inappropriate Use of "Probably Benign" in Mammography													x
165. Radiology – Nuclear Medicine Imaging Studies for Patients Undergoing Bone Scintigraphy													x
166. Rheumatoid Arthritis CMS Measures Group													x
167. Rheumatoid Arthritis – Disease Modifying Drug Therapy													x
168. Rheumatoid Arthritis - Tuberculosis Screen													x
169. Rheumatoid Arthritis – Periodic Assessment of Disease													x
170. Rheumatoid Arthritis – Functional Status Assessment													x
171. Rheumatoid Arthritis - Classification of Disease Prognosis													x
172. Rheumatoid Arthritis - Glucocorticoid Management													x
173. Safety Composite Assessmen t(ambulatory care)			x										
174. Screening – Blood Lead Level			x			x	x	x			x		
175. Screening – Breast Cancer (MN Community Measurement definition)				x			x	x	x		x		x
176. Screening – Evidence based Cervical Cancer				x			x		x		x		

Measure	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
177. Screening –Chlamydia	x					x		x	x	x	x	
178. Screening – Colorectal Cancer								x			x	x
179. Screening –Composite Cancer	x										x	
180. Screening – Preventative Services Composite for Adults				x								
181. Screening – Preventative Services Composite for Pediatrics				x								
182. Screening –Standardized Alcohol Abuse Screen	x			x								x
183. Screening – Tobacco Use												x
184. Screening – Tobacco Advising Smokers to Quit	x											x
185. Screening and Intervention – Alcohol Overuse Process for Emergency Department			x									
186. Spinal Surgery – Adverse Events	x											
187. Spinal Surgery - ODI	x											
188. Stroke and Stroke Rehab – CT or MRI Reports												x
189. Stroke and Stroke Rehab – Carotid Imaging Reports												x
190. Stroke and Stroke Rehab – DVT for Ischemic Stroke or Intracranial Hemorrhage												x
191. Stroke and Stroke Rehab – Discharge on Antiplatelet Therapy												x
192. Stroke and Stroke Rehab – Anticoagulant Therapy Prescribed for Atrial Fibrillation at Discharge												x
193. Stroke and Stroke Rehab – t-PA Considered												x
194. Stroke and Stroke Rehab – Screening for Dysphagia												x
195. Stroke and Stroke Rehab – Rehab Service Considered												x
196. Syncope												x
197. Thoracic Surgery – Recording of Clinical Stage for Lung Cancer and Esophageal Cancer Resection												x
198. Total knee arthroplasty – functional status measurement			x									
199. Total Hip or Total Knee Replacement											To be developed	
200. Upper Respiratory Infection –Appropriate Treatment for Children											x	x
201. Urinary Incontinence – Assessment												x
202. Urinary Incontinence – Characterization												x

203. Urinary Incontinence – Plan of Care											x
204. Well Child Visits – Infants							x		x	Retired 2007	
205. Well child visits -0-20 years old								x			
206. Well Child Visits – 3-6 Years Old							x			Retired 2007	
207. Well Child – Adolescents							x				
208. Wound care – Use of Compression in Venous Ulcers											x

‡‡

---

1) Hospitals - Medicare Hospital Compare Pay for Performance Program

The Reporting Hospital Quality Data for Annual Payment Update (RHQDAPU) program was originally mandated by Section 501(b) of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003. This section of the MMA authorized CMS to pay hospitals that successfully report designated quality measures a higher annual update to their payment rates. Initially, the MMA provided for a 0.4 percentage point reduction in the annual market basket (the measure of inflation in costs of goods and services used by hospitals in treating Medicare patients) update for hospitals that did not successfully report. The Deficit Reduction Act of 2005 increased that reduction to 2.0 percentage points.

In addition to giving hospitals a financial incentive to report the quality of their services, the hospital reporting program provides CMS with data to help consumers make more informed decisions about their health care. In FY 2007, nearly 95 percent of hospitals participated successfully in the reporting program and received the full market basket update for FY 2008. (Source: CMS, 2009)

The following is a list of Hospital Compare Measures.
[Acute myocardial infarction: percent of patients receiving fibrinolytic therapy during the hospital stay and having a time from hospital arrival to fibrinolysis of 30 minutes or less.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003549

[Acute myocardial infarction: percent of patients receiving PCI during the hospital stay with a time from hospital arrival to PCI of 90 minutes or less.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003551

[Acute myocardial infarction: percent of patients with a history of smoking cigarettes who receive smoking cessation advice or counseling during the hospital stay.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003545

[Acute myocardial infarction: percent of patients with LVSD and without both ACEI and ARB contraindications who are prescribed an ACEI or ARB at hospital discharge.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003544

[Acute myocardial infarction: percent of patients without aspirin contraindications who are](#)

[prescribed aspirin at hospital discharge.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003543

[Acute myocardial infarction: percent of patients without aspirin contraindications who received aspirin within 24 hours before or after hospital arrival.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003542

[Acute myocardial infarction: percent of patients without beta-blocker contraindications who are prescribed a beta-blocker at hospital discharge.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003546

[Acute myocardial infarction: percent of patients without beta-blocker contraindications who received a beta-blocker within 24 hours after hospital arrival.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003547

[Children's asthma care: percent of pediatric asthma inpatients who received relievers during hospitalization.](#) Joint Commission, The. 2008 Oct. NQMC:004378

[Children's asthma care: percent of pediatric asthma inpatients who received systemic corticosteroids during hospitalization.](#) Joint Commission, The. 2008 Oct. NQMC:004379

[Heart failure: percent of patients discharged home with written instructions or educational material given to patient or caregiver at discharge or during the hospital stay addressing all of the following: activity level, diet, discharge medications, follow-up appointment, weight monitoring, and what to do if symptoms worsen.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003555

[Heart failure: percent of patients with a history of smoking cigarettes, who are given smoking cessation advice or counseling during hospital stay.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003558

[Heart failure: percent of patients with documentation in the hospital record that LVS function was evaluated before arrival, during hospitalization, or is planned for after discharge.](#) Centers

---

for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003556

[Heart failure: percent of patients with LVSD and without both ACEI and ARB contraindications who are prescribed an ACEI or ARB at hospital discharge.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003557

[Pneumonia: percent of immunocompetent patients with community-acquired pneumonia who receive an initial antibiotic regimen during the first 24 hours that is consistent with current guidelines.](#) Centers for Medicare & Medicaid Services. 2007 Oct. NQMC:003566

[Pneumonia: percent of patients age 50 years and older, hospitalized during October, November, December, January, February, or March who were screened for influenza vaccine status and were vaccinated prior to discharge, if indicated.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003569

[Pneumonia: percent of patients aged 65 and older who were screened for pneumococcal vaccine status and were administered the vaccine prior to discharge, if indicated.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003560

[Pneumonia: percent of patients who had an assessment of arterial oxygenation by arterial blood gas measurement or pulse oximetry within 24 hours prior to or after arrival at the hospital.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003559

[Pneumonia: percent of patients who receive their first dose of antibiotics within 4 hours after arrival at the hospital.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct. NQMC:003565

[Pneumonia: percent of patients whose initial emergency department blood culture was performed prior to the administration of the first hospital dose of antibiotics.](#) Centers for Medicare & Medicaid Services/The Joint Commission. 2007 Oct

Appendix F: Ambulatory Measure Alignment Process

MMA completes a full inventory of P4P measures, including all required measures and pilot test measures. Inventory is sent to MN Community Measurement for review.

MNCM reviews required reporting measures for alignment across health plans.

Pilot test measures (individually negotiated with health plan, but not measured system-wide) are tracked for 1-2 years.

Measure is aligned.

Specs are approved by RAC and BOD and are listed in Standardized Measures catalog. Measure may or may not be publicly reported via MNCM.

Measure is not aligned.

Health plan brings pilot measure to MNCM for committee/community dialogue on whether to incorporate the measure into the Standardized Measures catalog.

Measure is discontinued.

MNCM works with interested parties to align measure.
Specs are approved by RAC and BOD and are listed in Standardized Measures catalog. Measure may or may not be publicly reported via MNCM.

Measure is not considered for alignment by MNCM.

MN Council of Health Plans then reviews the measure and Specs to determine next steps.

MNCM works with interested parties to align measure.
Specs are approved by RAC and BOD and are listed in Standardized Measures catalog. Measure may or may not be publicly reported via MNCM.

Measure is not included in Standardized measures.

MN Council of Health Plans then reviews the measure and Specs to determine next steps.

184